National Longitudinal Study of Adolescent Health

Wave III
Data Documentation

Carolina Population Center
University of North Carolina at Chapel Hill

April 2004

This research was supported by grant P01-HD31921 from the National Institute of Child Health and Human Development. Further information may be obtained by contacting Add Health, Carolina Population Center, 123 West Franklin Street, Room 403, Chapel Hill, NC 27516-2524 addhealth@unc.edu.

Wave III Data Documentation

Table of Contents

I. Introduction/Setup of Questionnaire

Identification of Respondent

Main Respondents

Wave III data contains follow-up interviews from 14,979 original Wave I respondents located by field interviewers between August 2001 and April 2002. The pretest data contain an additional 218 respondents who were interviewed during April 2001, for a Wave III total of 15,197. Interviews were conducted nationwide, including Alaska and Hawaii. Respondents who were overseas for the duration of the fieldwork or who were in the Armed Forces and deployed overseas for the duration of the fieldwork, were not eligible for a Wave III interview. Interviews were conducted mostly in home, but some interviews were conducted in school settings, workplaces, and other public places.

Interview laptops were preloaded with the respondent's name, gender, and birth date from earlier wave data, as well as the best address information that was available either from earlier wave contacts or from more recent tracing efforts conducted by the field contractor, RTI International. Respondents who could not be located on the basis of preloaded address information were located through tracing efforts conducted in the field by interviewers.

After the completion of fieldwork, cases that reported a name, gender, or birth date different from the preloaded information were evaluated to determine whether the correct person had been identified. In some instances, the evaluation used other earlier wave data to confirm past participation. Cases that could not be confirmed through this process were dropped, as their identity was suspect. In 13 instances it was determined that the interviews were conducted using the preloaded information for a sibling. These cases are corrected in the final file.

An IRB- and OHRP-approved prisoner protocol was developed to assist field interviewers to appropriately gain access to respondents who were located in correctional facilities during the fieldwork period. If incarcerated respondents were not expected to be released prior to the end of the fieldwork period, field interviewers were instructed to gain necessary permissions to conduct private and confidential interviews within the premises of the correctional facilities. The prisoner protocol stipulated that correctional facility administrators had to agree to the strict confidentiality requirements of Add Health.

Partners

Wave III also contains interviews for 1,507 romantic partners of Add Health respondents. The partner sample was a purposive, quota sample designed to collect information on one-third married, one-third cohabiting, and one-third dating partners. Eligibility criteria for this sample also stipulated that partners had to be opposite sex, 18 years old or older, and in a relationship with the main respondent lasting three months or longer. The recruitment of partners into this sample followed an indirect recruitment protocol that relied on the original respondent's assistance in passing along study contact information to the eligible partner.

Wave III Respondents

Sample Information

20,745	Wave I in-home respondents
+ 45	Wave II only genetic respondents (data for these respondents have never been released)

- 687	Wave I cases without a weight and without a genetic sample flag, not selected for Wave III
20,103	Wave III fielded sample

Interviewed Cases

15,170	Wave I respondents interviewed
+ 27	Wave II only genetic respondents interviewed
15,197	Wave III interviewed cases

Sample Weights

14,322	Wave III cases with weights
875	All Wave III cases without weights
27	Wave II only cases (need a Wave I weight to construct Wave III weights)
848	Wave I cases re-interviewed at Wave III that do not have Wave III weights

Informed Consent

To be eligible for a Wave III interview, original Add Health respondents were required to be 18 years of age or older. By the end of Wave III fieldwork, all original Wave III respondents were between 18 and 26 years old. The partner interview eligibility criteria also stipulated that the selected partner had to be 18 years old or older. There were no minors interviewed at Wave III, eliminating the need for parental consent. All Wave III respondents were asked to read and sign an informed consent form for participation in the Add Health survey. A copy of this form was retained by the respondent and the original copy was retained by the Field Interviewer (FI) for long-term storage. An electronic record indicating whether consent was provided was recorded in the laptop. All respondents who agreed to participate in the interview received an incentive payment of $20.

Administration of the AHPVT

At Wave III, the Add Health Picture Vocabulary Test (AHPVT), a computerized, abridged version of the Peabody Picture Vocabulary Test-Revised (PPVT-R), was re-administered. The AHPVT is a test of hearing vocabulary, designed for persons aged 2 1/2 to 40 years old who can see and hear reasonably well and who understand standard English to some degree.

The laptop computer was programmed to follow standard PPVT-R requirements and to appropriately skip items found to be too easy or too difficult for the respondent. The interviewer used the laptop in conjunction with the PPVT-R easel to administer the test. The interviewer was instructed to place the easel in plain view of the respondent while keeping the computer screen, which contained the stimulus word, out of view.

The interviewer read instructions from the computer screen, turned to the appropriate plate (page) in the AHPVT easel, pronounced the stimulus word, and entered the respondent's answer. Each test included a set of practice, or pretest items, followed by a series of test items arranged in order of increasing difficulty. The respondent was asked to listen to the word spoken by the interviewer and to select the picture on the plate that he or she believed best illustrated the meaning of the stimulus word. Once the response was entered into the computer, the program indicated the next plate to use in the test. In addition, the computer program determined test results automatically. These test results were not made available to the interviewer or to the respondent.

In an effort to minimize bias caused by variations in pronunciation of the stimulus words, each interviewer was required to pass a pronunciation certification for all words given in the AHPVT. The certification process was carried out as part of interviewer training. In addition, a pronunciation guide was provided as part of the computerized interview, as well as in the Field Interviewer Manual, and in the Showcard Booklet used by each interviewer.

Pretest and Main Study Differences: Questions That Were Added After the Pretest

Section A: Interview Setup

PRISON INTERVIEWER: IS THIS RESPONDENT TO BE INTERVIEWED AT A PRISON?

0 NO
1 YES

Section 16: Sexual Experiences and STDs

H3SE28 28. Have you ever been told that you have HIV or AIDS?

0 no
1 yes

Section 28: Tobacco, Alcohol, Drugs, Self-Image

H3TO26 26. How often do you smoke with your siblings?

1 very often
2 sometimes
3 never
4 I have no siblings

Section 13: Mentoring

H3MN19 19. Have you ever signed a pledge to abstain from sex until marriage?

1 yes
2 no *[skip to next section]*

20. In what month and year did you sign this pledge?

H3MN20M month _____

H3MN20Y year _____

[If H3MN20M and /or H3MN20Y = "don't know," ask Q.23.]

H3MN21 21. How old were you when you signed it?

age _____

Section 35: Interviewer's Report

H3IR20 20. Did the respondent require your help in completing the self-administered part of the questionnaire?

0 no
1 yes, all of the self-administered questions
2 yes, some of the self-administered questions

Preloaded Data and How It Guided Questionnaire Branching

Interviews with original Add Health respondents were pre-loaded with some Wave I and Wave II data, including the name, age, and sex of the respondent and identification of parent figures, friends, and siblings who appeared in earlier waves.

Gender: Main respondent's gender from earlier wave data; it was used to identify correct respondent and drive gender branching.

Wave III Gender

There are 20 cases in which Wave III gender (BIO_SEX3) does not match the Wave I gender (BIO_SEX). At Wave III, the preloaded gender variable came from the last wave of available data. Eighteen of these inconsistent cases match the Wave II gender (BIO_SEX2) and were confirmed at Wave III as being correct. Out of the remaining two inconsistent cases, there is one case where the Wave III gender, female, was confirmed by the Add Health security manager as being correct. In the last case, both the Wave I and Wave II gender is listed as male, which is correct. For this case only, the Wave III gender is incorrect.

Birth Date: Main respondent's birth date from earlier wave data; it was used to identify correct respondent.

PRMOM, PRPOP, PRYEAR: Main respondent's mother and father figures from earlier wave data. PRYEAR is the year of the last interview of the respondent, 1995 or 1996; it was used to drive branching patterns in Section 3.

Sibling questions, Section 5: The names of the respondent's other Add Health siblings, identified at earlier waves, were preloaded for the Wave III interview.

Criteria: Preload SIBLING = 0, 1, 2, 3, 4
If SIBLING = 0 then Section 5 is skipped.
If SIBLING = 1, 2, 3, or 4 then Section 5 is asked for each identified sibling (the SIBLING preload variable indicates the number of siblings).

Friend questions, Section 6: For respondents who were in the 7th or 8th grade at Wave I, an algorithm, based on clubs and activities from previous data, was used to select 10 names of students who also attended the same school. Wave III respondents, where FRIEND = 1, were asked to identify whether or not they were currently or had been previously friends with each of the 10 listed names.

Criteria: Preload FRIEND = 1
Respondents are asked questions in Section 6 about each of the 10 names with whom they indicate they are friends.

Sample flags were used to identify respondents who would be evaluated for special samples:

Partner Sample: Respondents selected for partner recruitment to fill quota samples of 500 married, 500 cohabiting, and 500 dating romantic partners.

Criteria: Preload PARTNER = 1
Criteria: Partner must be 18 or older, current partner, relationship has lasted three or more months, opposite sex partner.

Protocol required turning PARTNER to 0 once quota of partner interviews were filled.

Binge Drinking Sample: Respondents selected were asked to complete questions on attitudes toward binge drinking. All respondents with Wave I and II data who were in the 7th or 8th grade at Wave I were pre-identified. The questionnaire then evaluated the following criteria to determine if the binge drinking questions were asked.

Criteria: Preload BINGE = 1
Criteria, Group 1: Currently attending a two- or four-year college and never married.
Interview target of 350 males, 350 females.
Criteria, Group 2: NOT in college and never married.
Interview target of 350 males, 350 females.

Protocol required turning these selections off once these quotas were filled.

Partner interviews, of course, could not include pre-loaded data, since partners were new to the study. They were asked for name, age, and gender; questions about previous parent figures, friends, and siblings were not administered to partners.

Constructed Variables

Age: Constructed by subtracting the reported and confirmed date of birth from the interview date.

Pretest: The Add Health pretest, conducted during April 2001, used original respondents to test the Wave III questionnaire and procedures. These respondents are identified with the variable PRETEST. If PRETEST = 1, then the respondent participated in the pretest. Partners were not interviewed during the pretest.

Prison: Using several variables, PRISON, H3HR1, H3HR27, and interviewer comments, a single variable was used to indicate whether or not the Wave III interview was completed in a prison, jail, or detention center.

AHPVT Scores: The following AHPVT scores were constructed from the raw AHPVT score.

PVTPCT3—Wave III PVT percentile rank
PVTSTD3C—Wave III PVT standardized* score
PVTSTD3L—Wave III PVT standardized scores computed using the Wave I model for adolescents who were 18.5 years old to predict a standardized* score from the Wave III raw score.

*Raw scores were standardized to a mean of 100 and standard deviation of 15 before models were developed to predict smoothed scores from raw PVT scores.

II. Event History Calendar

Purpose

The Event History Calendar (EHC) was a special feature that was added to the Wave III questionnaire to improve accuracy in reporting lifetime event data. The EHC was used as a memory aid during both the CAPI and CASI portions of the interview to help respondents answer questions about events like romantic relationships, births and pregnancies, marriages, and graduations. Unlike most EHCs documented in the survey research literature, the Add Health EHC was not designed to be used as a data collection instrument. Rather, the interviewer and respondent worked together during the CAPI section to answer questions that would mark key events on the calendar. The respondent used it as a memory aid to provide a chronological framework when answering questions during the CAPI sections and a short CASI component.

Description

The EHC was organized into three domain columns: Public Events, where public landmark events were displayed; Personal Events, where personal landmark events were displayed; and Relationships, which displayed events involving the respondent's partners or spouses. The EHC displayed the respondent's age at the time of the public event, but the remainder of the information came from the respondent's answers to questions about important life events asked during the Wave III interview. For example, if the respondent reported a marriage, both the event and the date of that event were recorded in the EHC. The computer program compiled these life events in a database that was accessible to the respondent at key points in the interview. As the respondent continued to add life events in the computer database, those events appeared in date order on the calendar. From time to time, the respondent was asked to edit the calendar for accuracy. The EHC provided a visual chronicle representation of each respondent's important life events, allowing the researchers to obtain a more accurate depiction of those events.

III. Coding Schemes

Occupational Coding

The *1998 Standard Occupational Classification* (SOC) system was used by Add Health to classify respondent's first job and last or current job. The October 16, 2001, revisions were not available for use at the time of the development of the Add Health instrument and therefore were not included. Add Health respondents used this scheme to classify themselves into one of over 820 occupations by first selecting a job title from one of the 23 major groups, then progressing to select a job from one of the 96 minor groups, with the next selection coming from one of the 449 broad occupational categories, and the final job title coming from the list of 820 occupations. Additional information on the SOC is available at www.bls.gov/soc/.

Educational Majors

The *Classification of Instructional Programs: 2000 Edition* (CIP) was used by Add Health to code postsecondary fields of study in Section 7: Education. CIP codes replaced open-ended responses to questions 16, 17, 19, 20, 22, 29, and 30. According to the National Center for Educational Statistics web site (www.nces.ed.gov) the CIP: 2000 edition is the third revision, originally published in 1980 and revised in 1985 and 1990.

Ancestry

The ancestry variables H3OD7A, H3OD7B, H3OD7C, H3OD7D, and H3OD8 were recoded for cases of fewer than 10 respondents per country. Values for variables H3OD7A, H3OD7B, H3OD7C, and H3OD7D were combined to meet this criterion since respondents were not required to answer these questions in any particular order. This recoding was accomplished using the Geographic Subject Headings listing in Ovid Medline. A list of the same countries, taken from the parent source *National Library of Medicine's 2003 Medical Subject Headings* (MeSH), can be found at www.nlm.nih.gov/mesh/MBrowser.html using a search for geographic locations.

Religion

Christian S. Smith, UNC-CH Professor of Sociology, assisted Add Health in classifying and coding the responses to question 2 in Section 31: Religion and Spirituality. All respondents who answered "other" to question 1: "What is your present religion?" were asked to give us the name of the other religion. Question 2 contains the coded answers to this question.

IV. Biological Specimens and Physical Measurements

Sexually Transmitted Diseases

Specimen and Informed Consent

Add Health respondents were asked to provide a specimen of no more than 20 cc of first stream urine for Sexually Transmitted Disease (STD) testing. Respondents who agreed to provide a urine sample were asked to read and sign an informed consent form. Respondents received an incentive payment of $10 for providing this sample. Urine samples were collected in the field and mailed via Federal Express in refrigerated containers to UNC Laboratories where they were assayed for chlamydia and gonorrhea.

Collection Process

The collection, handling, shipping, and FI training protocols were carefully developed to ensure that the urine samples would arrive on time and in the correct condition to be analyzed and reported to the respondent. Shipments of urine samples had special requirements. The urine had to be kept cold, but not frozen, after collection and until its arrival at the labs. If the urine arrived at the labs warm, the results could not be supplied to the respondent. And there was a strict 96-hour limit on the time between the collection and the time of arrival to the labs. If a sample arrived that did not meet this criterion then the test results were not available to be reported to the respondent. In addition, respondents could not have urinated within an hour of the time of the urine collection procedures, so the FIs were taught to be vigilant, obtain verbal consent, and instruct the respondent how to collect urine if they headed off to the bathroom during the interview but prior to the usual collection time.

Tests Performed

The urine specimens were analyzed for the presence of three STDs: chlamydia, gonorrhea, and trichomoniasis. The results were reported as positive or negative for the presence of each STD.

Chlamydia C. trachomatis was diagnosed with a molecular diagnostic test called the ligase chain reaction. This assay tests for the presence of DNA from *C. trachomatis*. The assay is manufactured by Abbott Laboratories, Inc. The test is FDA approved for the detection of *C. trachomatis* DNA in male and female specimens, male urethral swabs, and female endocervical swabs.

This assay has a reported specificity ranging from 87.1 to 100% when testing urine samples according to the manufacturer's package insert. A one milliliter aliquot of urine was centrifuged, the supernatant aspirated, and the pellet resuspended in urine resuspension buffer. The re-suspended pellet was then heated at 97°C for 15 minutes. After cooling to room temperature, 100ul of the processed sample was added to the LCx Chlamydia amplification vial and placed in the thermal cycler. After amplification, the amplification vials (patient samples, controls, and calibrators) were placed into the LCx analyzer. The detection run was then automatically completed and the results reviewed and entered into the computer database. A positive result was defined as a sample with a signal to cutoff ratio of $\geq$1.00. An indeterminate result was not possible as any initially equivocal result (signal to cutoff 0.80 – 0.99) was repeated to determine the final result ($\geq$1.00 = positive, <1.00 = negative). Any sample that yielded a signal to cutoff ratio of $\geq$0.80 was repeated. This included all samples with initial signal to cutoff result $\geq$1.00.

The laboratory results provided include positive, negative, or not tested. A positive result indicates that *C. trachomatis* DNA was detected in the urine of the respondent. Since the specificity of the test is not 100% it is possible that a result may be falsely positive. A negative result indicated that *C. trachomatis* DNA was not detected in the urine of the respondent. Since the test is not 100% sensitive, it is possible that the result could be falsely negative. A sample-not-tested result indicates that for a variety of potential reasons, the sample was either not received, received in a state not suitable for testing, or not tested due to a laboratory error.

Gonorrhea *N. gonorrhea* was diagnosed with a molecular diagnostic test called the ligase chain reaction. This assay tests for the presence of DNA from *N. gonorrhea*. The assay is manufactured by Abbott Laboratories, Inc. The test is FDA approved for the detection of *N. gonorrhea* DNA in male and female urine specimens, male urethral swabs, and female endocervical swabs.

This assay has a reported specificity ranging from 92.2 to 100% when testing urine samples according to the manufacturer's package insert. A one milliliter aliquot of urine was centrifuged, the supernatant aspirated, and the pellet resuspended in urine resuspension buffer. The re-suspended pellet was then heated at 97°C for 15 minutes. After cooling to room temperature, 100ul of the processed sample was added to the LCx Neisseria amplification vial and placed in the thermal cycler. After amplification, the amplification vials (patient samples, controls and calibrators) were placed into the LCx analyzer. The detection run was then automatically completed and the results reviewed and entered into the computer database. A positive result was defined as a sample with a signal to cutoff ratio of $\geq$1.00. An indeterminate result was not possible as any initially equivocal result (signal to cutoff 0.80 – 0.99) was repeated to determine the final result ($\geq$1.00 = positive, <1.00 = negative). Any sample that yielded a signal to cutoff ratio of 0.80 – 0.99 was repeated.

The laboratory results provided include positive, negative, or not tested. A positive result indicates that *Neisseria gonorrhea* DNA was detected in the urine of the respondent. Since the specificity of the test is not 100% it is possible that a result may be falsely positive. A negative result indicated that *Neisseria gonorrhea* DNA was not detected in the urine of the respondent. Since the test is not 100% sensitive, it is possible that the result could be falsely negative. A sample-not-tested result indicates that for a variety of potential reasons, the sample was not received, received in a state not suitable for testing, or not tested due to a laboratory error.

Trichomonas Vaginalis Urine specimens were tested using a PCR-ELISA that detects *Trichomonas vaginalis* DNA. The *T. vaginalis* PCR-ELISA was developed at the STD Cooperative Research Center's Microbiology Core Laboratory at the University of North Carolina at Chapel Hill. Formal validation studies of the performance of the assay compared to wet mount and culture are published in the *Journal of Clinical Microbiology* (Kaydos et al., 2002, 40:89-95 and Kaydos-Daniels et al., 2003, 41:318-323). The PCR primers used in this assay, TVK3 and TVK7, specifically amplify a 312-bp sequence from repetitive DNA in the *T. vaginalis* genome. PCR with these primers is negative with human DNA, other organisms found in the genitourinary tract, and other *Trichomonas* species (Kengne et al., 1994, *Cellular and Molecular Biology* 40:819-831). PCR products were detected using an ELISA format with a nucleotide probe (TVK) that binds a specific DNA sequence in the amplified product from positive specimens. Results were determined from absorbance values that ranged from 0 to 3.5. Based on the published ROC analysis of the test in urine from women, an ELISA absorbance cutoff of 3.0 was used to determine a positive result for specimens from female subjects. The published ROC analysis of the test in urine from men established 2.0 as the cutoff for male subjects. Negative test results were coded 0, and positive test results were coded 1.

The assay procedures were modified slightly from the published references cited above; Add Health specimens were tested using the specifics detailed herein. Urine specimens were received and logged into the computer system in the UNC Hospitals Microbiology Laboratory. A 1mL aliquot of each urine sample was labeled with the UID number, and samples were transported to the STD Cooperative Research Center Microbiology Core Laboratory for *T. vaginalis* testing. Samples were prepared within two days of receipt using the Amplicor CT/NG Urine Specimen Prep Kit (Roche Diagnostic Systems, Indianapolis, IN) according to the manufacturer's instructions. Urine preps were stored at -70°C until PCR-ELISA testing.

For PCR, 50 mL of thawed urine prep was used as template in a reaction containing 40 pmol each of primers TVK3 and TVK7 (digoxigenin [DIG] labeled), 200 mM each dATP, dCTP, dGTP, and dUTP, 2.5 U *Taq* polymerase (Gibco BRL, Grand Island, NY) 4mM $MgCl_2$ and 1.0 U AmpErase (uracil N-glycosylase; Applied Biosystems, Foster City, CA) in 1x PCR buffer (Gibco BRL) in a final volume of 0.1 mL. PCR consisted of an initial five-minute incubation at 95°C followed by 40 cycles of denaturation at 90°C for one minute, annealing at 60°C for 30 seconds, and extension at 72°C for two minutes. Purified *T. vaginalis* DNA and sterile water were used as positive and negative controls with every batch of PCR tests. Unused portions of the urine preps are being held in -70°C storage in the Microbiology Core Laboratory.

PCR products were detected using the PCR ELISA DIG detection kit (Roche Diagnostic Systems) according to the manufacturer's instructions (included with this document). For ELISA, 60 mL of PCR product was incubated with 50 ng of biotinylated TVK probe in hybridization buffer in a final volume of 1.0mL; 200 mL of the hybridization solution was pipetted into streptavidin-coated wells of the kit microtiter plate and incubated at 37°C for two hours with gentle rocking. Plates were washed five times in a microtiter plate washer containing kit washing solution, and 200 mL of anti-DIG POD solution (diluted as per kit instructions) was added to each well and incubated at 37°C for 30 minutes with gentle rocking. Plates were washed as above, and 200 mL of ABTS substrate solution was added to each well and incubated at 37°C in the dark for one hour with gentle rocking. Absorbance was read at 405nm in a spectrophotometric plate reader with a linear range up to 3.5; absorbance of the kit negative control read at 492 nm was subtracted from sample readings as directed in the kit instructions. Tests were considered valid if the absorbance of the positive control was $\geq$2.0 and the negative control was <0.2. Duplicate values for individual specimens were averaged, and absorbance values were recorded in an Excel database maintained in the Microbiology Core Laboratory. Based on the published ROC analyses of the test in urine, ELISA absorbance values (Abs) of 3.0 for specimens from women and 2.0 for specimens from men were used as cutoffs. For GENDER = 1, Abs <2.0 was considered negative and coded as 0 in the Tvaginalis.xls results file; GENDER = 1, Abs $\geq$2.0 was considered positive and coded as 1. For GENDER = 2, Abs <3.0 was considered negative and coded as 0 in the results file; GENDER = 2, Abs $\geq$3.0 was considered positive and coded as 1. Indeterminate results were not possible, and positive results were not verified by retesting.

T. vaginalis PCR-ELISA is not FDA approved. It is currently used for research purposes only.

Height Measurement

An exact height measurement was taken for all respondents who agreed to allow the field interviewer to take this measurement. The height measurement was taken to the nearest 1/8 inch. Field interviewers were instructed to stand on a chair to take this measurement if they needed to in order to be taller than the respondent. The respondent was instructed to take off his or her shoes and stand up against a wall, preferably on a non-carpeted surface. The respondent was asked to stand up straight with eyes facing forward. The FI used a level to measure the exact height of the respondent by marking the height with a small self-sticking

paper on the wall. The FI then set the measuring tape against the wall and measured from the floor to the top edge of the paper. The FI then wrote the exact height to the nearest 1/8 inch on the paper, and then transferred the feet, inches, and partial inches information into the appropriate fields on the computer screen.

Weight Measurement

A weight measurement was taken for all respondents who agreed. The weight measurement was taken using a digital bathroom scale that measured weight to the nearest 1/2 pound. The field interviewer was asked to position the scale on a flat, even surface that would provide sufficient space to take a reading while the respondent was still on the scale. The respondent was asked to take off his or her shoes and stand on the scale. If the respondent indicated that he or she wanted to take off extra layers of clothing then that was allowed. The FI was instructed to wait for the scale to settle and read the scale while the respondent was still standing on it. The FI then recorded the weight to the nearest half (1/2) pound in the appropriate field on the computer screen. If the respondent weighed more than 330 pounds, the scale maximum, the FI entered a special code, 888, indicating the respondent's weight was over the scale maximum.

V. Wave III Weights Calculation

(Developed by Paul P. Biemer and Elvessa D. Aragon, RTI International, Research Triangle Park, NC)

Introduction

In this chapter, we describe the process for weighting the Wave III data to compensate for Wave III nonresponse. Although alternative methods for compensating for Wave III nonresponse were considered, the approach described here is essentially the same approach that was used for Wave II, which was also consistent with the Wave I nonresponse approach. Thus, we begin with a brief review of the weighting approaches for Wave I and Wave II and then describe our approach for Wave III.

Review of Weights from Wave I and Wave II

Wave I Grand Sample Weight

In Wave I, the weight for the grand sample was calculated using a standard multiplicity approach. That is, the weights for each case are summed across the various samples for which a sample case was selected, and then this sum is divided by the number of samples for which the case was eligible. The nonresponse adjustments for both the school level and student level as well as post-stratification adjustments from each of the samples are automatically incorporated into this grand sample weight. The grand sample Wave I weight for case *j*, denoted by W_{1j}, is given by

$$W_{1j} = \frac{\sum W_{ij}}{s_j} \qquad (1)$$

where s_j is the number of samples for which case *j* was eligible and W_{ij} is the weight of case *j* for sample *I*.

Nonresponse adjustments were recalculated for three of the genetic supplements (twins, full siblings, and half siblings) before they were used in Equation (1). Originally, the nonresponse adjustments for these samples were based on the pair-level nonresponse. The new nonresponse adjustment is based on the response rates of the individuals. The sample weights for these three genetic supplements were recalculated using the individual-level nonresponse adjustment before they were included in the grand sample weight.

The grand sample weight was trimmed to eliminate weights greater than 6,000, and poststratified to 1995 Current Population Survey estimates of the size of each grade-sex-race (black vs. non-black) subpopulation.

Wave II Grand Sample Weight

The Wave II sample comprised primarily adolescents who participated in Wave I, with the following changes:

- Wave I 12th graders were not retained, except the adolescents who were part of genetic pairs
- Disabled and siblings of twins were not retained
- Some adolescents were reclassified into groups different from the groups for which they were selected in Wave I. However, the Wave I sample classifications were used for weighting purposes.

The grand sample weight for Wave II was derived from the Wave I grand sample weight with a Wave II nonresponse adjustment calculated separately for each sample school. The Wave II grand sample weight for case *j* is given by

$$W_{2j} = \frac{W_{1j}}{R_2} \qquad (2)$$

where R_2 is the school-specific Wave II (weighted) response rate calculated using the Wave I grand sample weight W_{1j}. It is not explicitly stated in the weighting documentation for the previous waves whether the Wave II grand sample weight was trimmed in the same manner as the Wave I grand sample weight. The Wave II grand sample weight was poststratified to the 1995 population, except for the few Wave I 12th graders who were retained in the Wave II sample. The Wave II weights for these 12th graders were not inflated to reflect the 1995 population of 12th graders in the United States.

Calculation of Weights for Wave III

The same basic approach for adjusting the Wave I weight for Wave II nonresponse was used to adjust the Wave I and Wave II weights for Wave III nonresponse. These weight adjustments were applied to form cross-sectional and longitudinal analysis weights for both the grand sample and the public-use sample. Following the same weight adjustment approach for all three waves should minimize the impact of the weight adjustment methodologies at each wave on data analysis.

Wave III of Add Health consists of a follow-up interview with the 20,745 in-home sample members from Wave I. Interviews were conducted between August 2001 and April 2002. In addition, the sample fielded for Wave III included 45 respondents who were interviewed at Wave II only, in order to increase the number of respondents in the genetic sample. These 45 cases were never assigned weights in Wave II and the data from Wave II have never been cleaned or used. Excluded from the Wave III sample were 687 cases that did not have sample flags or weights from Wave I. Thus, the total number of cases fielded for Wave III is 20,103.

Eligibility for Wave III

The 687 cases that were not fielded in Wave III are ineligible and were not assigned Wave III weights. Similarly, the 45 cases added in Wave II were never assigned a weight in Wave II and the data from Wave II have never been cleaned or used. For purposes of weighting, these 45 cases were considered ineligible and were not be assigned any weights for Wave III.

Of all the groups of cases that would usually be deemed ineligible, it was determined that "deceased" is the only group that has the same definition at Wave II and Wave III. The other groups (non-existent, permanently out of the country, ineligible due to age, not a sibling of the originally sampled adolescent, or on active military duty) have different meanings at Wave II and Wave III. Hence, they are not comparable. Thus, only the 96 deceased cases will be further deemed ineligible for purposes of Wave III weighting. These cases were not assigned Wave III weights. Table 1 shows the final disposition of the Wave III cases.

Table 1. Final Disposition of Wave III Cases

KEEPCASE	Wave III Final Disposition		Wave I	Wave II	Wave III	Freq	Eligible*	Respondent*
Excluded from Wave III because these cases have no sample flags nor Wave I weights								
0	!		1	!	!	538	I	–
0	!		1	1	!	149	I	–
					Total	687		
Cases found to be Wave III ineligible (for weighting purposes)								
1	459	Deceased	1	!	!	40	I	–
1	459	Deceased	1	1	!	56	I	–
					Total	96		
Cases added in Wave II to supplement the genetic sample; these have no Wave I or II data nor weights								
1	422	Unavailable after repeated attempts	!	!	!	2	I	
1	458	R name not identified	!	!	!	2	I	
1	460	Final refusal by R	!	!	!	5	I	
1	477	Incarcerated	!	!	!	1	I	
1	481	Moved out of interviewing area: No FI available	!	!	!	1	I	
1	484	Not located: leads exhausted	!	!	!	4	I	
1	489	Other non-interview (specify)	!	!	!	3	I	
1	494	Interview completed (all sections)	!	!	1	27	I	
					Total	45		
Eligible, but non-interviews in Wave III								
Cases considered eligible for weighting purposes								
1	451	R age ineligible	1	!	!	1	E	NR
1	451	R age ineligible	1	1	!	3	E	NR
1	482	Moved out of country	1	!	!	66	E	NR
1	482	Moved out of country	1	1	!	109	E	NR
1	486	Active duty military unavailable for duration	1	!	!	46	E	NR

KEEPCASE	Wave III Final Disposition		Wave I	Wave II	Wave III	Freq	Eligible*	Respondent*
1	486	Active duty military unavailable for duration	1	1	!	141	E	NR
Cases where person interviewed was not the correct Add Health respondent								
1	494	Interview completed (all sections) *but WAVE3* = !	1	!	!	13	E	NR
1	494	Interview completed (all sections) *but WAVE3* = !	1	1	!	25	E	NR
Other eligible nonrespondents								
1	420	No one home after repeated attempts	1	!	!	22	E	NR
1	420	No one home after repeated attempts	1	1	!	17	E	NR
1	422	Unavailable after repeated attempts	1	!	!	239	E	NR
1	422	Unavailable after repeated attempts	1	1	!	419	E	NR
1	423	Unavailable for duration of field study	1	!	!	32	E	NR
1	423	Unavailable for duration of field study	1	1	!	58	E	NR
1	458	R name not identified	1	!	!	4	E	NR
1	458	R name not identified	1	1	!	9	E	NR
1	460	Final refusal by R	1	!	!	421	E	NR
1	460	Final refusal by R	1	1	!	591	E	NR
1	461	Final refusal by parent/guardian	1	!	!	43	E	NR
1	461	Final refusal by parent/guardian	1	1	!	75	E	NR
1	462	Final refusal by other	1	!	!	13	E	NR
1	462	Final refusal by other	1	1	!	17	E	NR
1	470	Language barrier - Spanish	1	!	!	4	E	NR
1	470	Language barrier - Spanish	1	1	!	3	E	NR

KEEPCASE	Wave III Final Disposition		Wave I	Wave II	Wave III	Freq	Eligible*	Respondent*
1	471	Language barrier - other	1	1	!	2	E	NR
1	475	Physically/mentally incapable	1	!	!	18	E	NR
1	475	Physically/mentally incapable	1	1	!	23	E	NR
1	477	Incarcerated	1	!	!	32	E	NR
1	477	Incarcerated	1	1	!	71	E	NR
1	478	Institutionalized	1	1	!	3	E	NR
1	481	Moved out of interviewing area: no FI available	1	!	!	9	E	NR
1	481	Moved out of interviewing area: no FI available	1	1	!	18	E	NR
1	483	Not located by end of field period: leads not exhausted	1	!	!	39	E	NR
1	483	Not located by end of field period: leads not exhausted	1	1	!	51	E	NR
1	484	Not located: leads exhausted	1	!	!	872	E	NR
1	484	Not located: leads exhausted	1	1	!	1,253	E	NR
1	485	Mistakenly interviewed as partner	1	!	!	1	E	NR
1	485	Mistakenly interviewed as partner	1	1	!	2	E	NR
1	489	Other non-interview (specify)	1	!	!	5	E	NR
1	489	Other non-interview (specify)	1	1	!	13	E	NR
1	501	Interview completed but data lost *but WAVE3* = !	1	1	!	2	E	NR
1	502	Unknown outcome; ID number used to interview another respondent, outcome for this case not recorded	1	!	!	2	E	NR

KEEPCASE	Wave III Final Disposition		Wave I	Wave II	Wave III	Freq	Eligible*	Respondent*
1	502	Unknown outcome; ID number used to interview another respondent, outcome for this case not recorded	1	1	!	5	E	NR
					Total	4,792		
Eligible interviews								
1	490	Breakoff/partial Interview	1	!	1	2	E	R
1	490	Breakoff/partial Interview	1	1	1	1	E	R
1	491	Interview completed only	1	1	1	4	E	R
1	492	Interview completed with FI observations	1	1	1	16	E	R
1	493	Interview completed with GPS	1	!	1	1	E	R
1	493	Interview completed with GPS	1	1	1	10	E	R
1	494	Interview completed (all sections)	1	!	1	3,546	E	R
1	494	Interview completed (all sections)	1	1	1	11,590	E	R
					Total	15,170		
					Grand Total	20,790		

* E = Eligible; I = Ineligible; R = Respondent; NR = Non-respondent

Weights for the Grand Sample

Because the Wave III sample consists of Wave I in-home sample members, creating the Wave III **cross-sectional weight** for the grand sample simply involved adjusting the Wave I grand sample weight for additional Wave III nonresponse. The basic formulas for these adjustments appear in Appendix A.

Weighted nonresponse adjustments were calculated separately for each school using the Wave I grand sample weight. Wave III respondents who also have a Wave I grand sample weight (i.e., were Wave I respondents) were assigned a cross-sectional weight. Weights were missing, otherwise.

The nonresponse-adjusted cross-sectional weight was poststratified to estimates of the grade-sex-race subpopulations derived from the Wave I grand sample weight, excluding the deceased at Wave III. These estimates reflect the portion of the 1995 population (represented by the Wave I sample) that would have been eligible at Wave III. The estimates were calculated by summing the Wave I grand sample weight or all the sample members of each grade-sex-race domain that were alive at Wave III. Tables 2 and 3 show the grade-sex-race subpopulation estimates used to poststratify the cross-sectional weight to the Wave I population.

Table 2. Wave I Grand Sample Weight for All Cases Alive at Wave III
(Wave I subpopulation estimates for Wave III poststratification domains)

Gender	Grade	All Races	Black	Non-Black
Male	6	81	!	81
	7	1,941,288	278,123	1,663,165
	8	1,901,959	307,223	1,594,735
	9	1,937,867	312,653	1,625,214
	10	1,723,584	312,676	1,410,908
	11	1,775,968	269,069	1,506,899
	12	1,959,261	337,170	1,622,091
	Total	11,240,009	1,816,915	9,423,094
Female	!	2,168	!	2,168
	7	1,922,258	295,601	1,626,657
	8	1,725,692	295,998	1,429,694
	9	1,828,309	276,474	1,551,835
	10	1,856,561	339,561	1,517,000
	11	1,654,368	289,608	1,364,760
	12	1,876,974	324,280	1,552,693
	Total	10,866,331	1,821,523	9,044,808
Total		22,106,340	3,638,438	18,467,902

Table 3. Wave I Public-Use Sample Weight for All Cases Alive at Wave III
(Wave I subpopulation estimates for Wave III poststratification domains)

Gender	Grade	All Races	Black	Non-Black
Male	7	1,922,013	266,402	1,655,611
	8	1,848,950	310,077	1,538,873
	9	2,011,608	331,846	1,679,762
	10	1,777,445	340,111	1,437,334
	11	1,741,846	252,734	1,489,112
	12	1,941,619	329,796	1,611,823
	Total	11,243,480	1,830,965	9,412,515
Female	7	1,910,517	302,111	1,608,407
	8	1,737,703	282,166	1,455,537
	9	1,862,749	290,479	1,572,270
	10	1,874,094	319,861	1,554,233
	11	1,660,018	302,717	1,357,301
	12	1,855,765	321,436	1,534,328
	Total	10,900,845	1,818,769	9,082,076
Total		22,144,325	3,649,734	18,494,591

The Wave III **longitudinal weight** was calculated similarly, except that the Wave II grand sample weight was used instead of the Wave I grand sample weight (see Basic Weighting Formulas for Wave III in Appendix A). That is, the Wave II grand sample weight was adjusted for additional Wave III nonresponse using weighted nonresponse adjustments to the Wave II grand sample weight calculated separately for each school. Sample members who were respondents in both Wave II and Wave III were assigned weights. Weights were missing, otherwise. Wave II respondents were defined as sample members who had a Wave II grand sample weight. For purposes of this weighting process, Wave II ineligibles were those who were deceased, those who were not selected for Wave II interview, and the 45 cases added in Wave II but never weighted.

The Wave II weight of the few Wave I 12th graders who were retained in Wave II was not poststratified to reflect the corresponding population of 12th graders in the United States. In order to keep the number represented by these seniors at Wave II consistent with estimates produced by the Wave II sample, estimates of the grade-sex-race subpopulations derived from the Wave II grand sample weight, excluding the deceased, were used to poststratify the nonresponse-adjusted longitudinal weight. The estimates were calculated by summing the Wave II grand sample weight for all the sample members of each grade-sex-race domain that were still alive at Wave III. Tables 4 and 5 show the subpopulation estimates used to poststratify the longitudinal weight to the Wave II population.

Table 4. Wave II Grand Sample Weight for All Cases Alive at Wave III
(Wave II subpopulation estimates for Wave III poststratification domains)

Gender	Grade	All Races	Black	Non-Black
Male	7	1,868,725	260,137	1,608,588
	8	1,865,609	298,846	1,566,764
	9	1,900,654	295,740	1,604,914
	10	1,702,816	304,138	1,398,678
	11	1,690,578	248,973	1,441,605
	12	389,685	65,396	324,289
	Total	9,418,067	1,473,229	7,944,838
Female	7	1,907,963	282,017	1,625,946
	8	1,730,370	294,073	1,436,297
	9	1,832,326	271,139	1,561,186
	10	1,863,175	343,823	1,519,352
	11	1,614,018	277,480	1,336,538
	12	388,528	62,030	326,499
	Total	9,336,381	1,530,562	7,805,819
Total		18,754,449	3,003,791	15,750,657

Table 5. Wave II Public-Sse Sample Weight for All Cases Alive at Wave III
(Wave II subpopulation estimates for Wave III poststratification domains)

Gender	Grade	All Races	Black	Non-Black
Male	7	1,897,714	250,504	1,647,210
	8	1,847,449	303,767	1,543,682
	9	1,985,721	316,569	1,669,151
	10	1,789,051	330,115	1,458,936
	11	1,688,328	241,534	1,446,794
	12	288,934	65,733	223,201
	Total	9,497,196	1,508,223	7,988,973
Female	7	1,902,886	282,139	1,620,747
	8	1,719,583	272,313	1,447,270
	9	1,829,751	276,654	1,553,097
	10	1,874,247	335,928	1,538,318
	11	1,648,270	302,560	1,345,709
	12	258,365	45,130	213,235
	Total	9,233,102	1,514,725	7,718,377
Total		18,730,298	3,022,948	15,707,350

Weight Trimming

Since the nonresponse-adjusted weights were trimmed in the previous two waves, the question of whether to trim the weights in Wave III was considered. Weight trimming can reduce the variances of estimates but may also increase their biases. The hope is that trimming will reduce the total mean squared error, but whether or not this is accomplished cannot be evaluated with the available data. Nevertheless, a good reason to trim the weights for Wave III is to be consistent with the previous two waves.

To evaluate the effect of trimming on the nonresponse-adjusted weights on the precision of the estimates, we calculated the unequal weighting effects (UWE) for the cross-sectional and longitudinal weights both with and without trimming. The UWE is defined as $[1+ CV^2(\text{weights})]$ where $CV^2(\text{weights})$ refers to the coefficient of variation of the weights squared. As was done in Wave I, the trimmed value was set to 6000; i.e., all nonresponse-adjusted weights greater than 6000 were set equal to 6000. Poststratification adjustments were then applied to both the untrimmed and the trimmed nonresponse-adjusted weights. Table 6 summarizes the results of our analysis of the trimming effects. Note that the public-use sample weight was not trimmed in previous waves, so it will not be trimmed in Wave III.

Table 6. Effect on Weight Trimming on the UWE for Cross-Sectional and Longitudinal Weighting

Type of Weight	Number of Positive Weights	Untrimmed Weights UWE	Trimmed Weights	
			Number Trimmed	UWE
Cross-sectional				
Adjusted for NR	14,322	1.92370	209	1.82769
Final (Poststratified)	14,322	1.93530	209	1.84260
Longitudinal				
Adjusted for NR	10,828	1.86821	209	1.74612
Final (Poststratified)	10,828	1.87360	209	1.75417
Wave I Final	18,924			1.88615
Wave II Final	13,568			1.82098

From Table 6, note that the UWE for the untrimmed Wave III longitudinal weight is approximately at the level of the final weight for Wave I (i.e., about 1.9 in both cases). Thus, trimming does not appear to be necessary for Wave III longitudinal weighting to reduce the UWE to that of Wave I. For the cross-sectional weight, the UWE for the poststratified untrimmed weight is only slightly higher than the Wave I UWE. These results suggest that the UWE is only slightly reduced by trimming at the risk of an unknown increase in nonsampling bias. Further, the untrimmed UWE is close to that of Wave I in both cases, suggesting that the untrimmed UWE is acceptable for both longitudinal and cross-sectional weighting. Nevertheless, since the percentage of weights that are trimmed is small (between 1.5 and 1.9 percent) the risk of nonsampling bias by trimming is also small. Therefore, we decided to trim the weights to be consistent with the treatment of the weights in the earlier waves. We recommend that the trimmed final weights be used in analyses with Wave III data.

VI. Questionnaire Sections

Section Number	Section Name	Variable Root Name
1	Overview and Demographics	H3ODxx
2	Household Roster and Residence History	H3HRxx
3	Parental Support and Relationships	H3WPxx
4	Retrospective ADHD	H3RAxx
5	Relationships with Siblings	H3WSxx
6	Friends	H3FSxx
7	Education	H3EDxx
8	Labor Market Experience and Active-Duty Military Service	H3LMxx
9	General Health and Diet	H3GHxx
10	Access to Health Services, Health Insurance	H3HSxx
11	Illnesses, Medications, and Physical Disabilities	H3IDxx
12	Social Psychology and Mental Health	H3SPxx
13	Mentoring	H3MNxx
14	Marriage/Co-habitation History and Attitudes	H3MRxx
15	Economics and Personal Future	H3ECxx
16	Sexual Experiences and STDs	H3SExx
17	Table of Relationships	H3TRxx
18	Table of Pregnancies	H3TPxx
19	Relationships in Detail	H3RDxx
20	BEM Inventory	H3BMxx
21	Propensity for Risk	H3PRxx
22	Completed Pregnancies	H3PGxx
23	Current Pregnancies	H3PCxx
24	Live Births	H3LBxx
25	Children and Parenting	H3KKxx
26	Delinquency and Violence	H3DSxx
27	Involvement with the Criminal Justice System	H3CJxx
28	Tobacco, Alcohol, Drugs, Self-Image	H3TOxx
29	Mistreatment by Adults	H3MAxx
30	Civic Participation and Citizenship	H3CCxx
31	Religion and Spirituality	H3RExx
32	Gambling	H3GMxx
33	Daily Activities	H3DAxx
34	Biological Specimens	H3BPxx
35	Interviewer Report	H3IRxx

VII. Data Cleaning

The following procedures were applied when cleaning the Wave III interview data.

A. When the root question was answered, the following cleaning procedure was used.

1. Use the root question to determine how to code responses to sub-questions.
2. If the sub-questions should have been skipped because the root question implied a skip, then all sub-questions should be coded as legitimate skips, regardless of data received.
3. All other combinations of answers to the sub-questions where the root question indicates that the sub-questions should be answered would be left as received.

Exceptions

1. When a programmer felt that the above rule was not applicable, then a decision on how to recode the question was made based on the data for the respondent.
2. These cases are documented at the end of the questionnaire section in a way that explains how and why the question was recoded.

B. When the root question is missing, use the following rule.

1. Infer the answer to the root question, based on the answers to the sub-questions, if the sub-question answers are unambiguous. Use the following guidelines.
   a. If the root is not answered and none of the sub-questions are answered, code the root question to correspond with the branching logic and code the sub-questions as legitimate skips.
   b. Infer the root to be the non-branching answer if *any* of the sub-questions are answered, regardless of the type of responses given to the sub-questions.

C. Data type specific recoding, cleaning for Wave III.

1. Data were evaluated within sections for inconsistencies. Rules were established for all data changes so that changes were applied in a consistent manner across respondents. Changes to the data are documented in separate notes at the end of the codebook. The reasons for changing the inconsistent data are based on the responses to questions in other parts of the interview and cannot be generally specified.
2. When records are deleted from “offshoot” data files, the variable that controlled the creation of the number of records in the “offshoot” file was adjusted to reflect the new number of records.
3. Rules were established for recoding individual items such as four-digit year dates, height, and weight. For example, values of 98, 99, 95 were changed to 1998, 1999, 1995 and values of 1, 2, or 0 were recoded to missing. Other types of out of range responses were recoded to missing.

D. How to assign reserve and consistency codes.

1. Respondents were allowed to give a response of don’t know, refused, or not applicable to any question. The following reserve codes are used throughout the Wave III data.

6, 96, 996, 9996 = refused

8, 98, 998, 9998 = don’t know

9, 99, 999, 9999 = not applicable

2. In addition, the following consistency codes are used.

   5, 95, 995, 9995 = question not asked because the respondent is not part of the sample selected to receive the question, for example, sibling, friend, binge.

   7, 97, 997, 9997 = legitimate skip

VIII. Appendices

Appendix A: Basic Weighting Formulas for Wave III

Define the following symbols:

D the domain of interest; for example, a particular grade-sex-race cell

S the original Wave I sample

R_{Dw} the set of respondents at Wave *w*, for *w* = 1,2,3

E_{Dw} the set of eligible persons at Wave *w*, for *w* = 1,2,3

I_{Dw} the set of ineligible persons at Wave *w*, for *w* = 1,2,3

ω_{1i} Wave I final weight (either grand sample or public-use sample, depending upon the context) defined for all $I \in R_{D1}$

ω_{2i} Wave II final weight (either grand sample or public-use sample, depending upon the context) defined for all $I \in R_{D2}$

Grand Sample Cross-Sectional Weight

For each sample school, define the school-specific weighted response rate for Wave III given Wave I response as

$$r_{D3|1} = \frac{\sum_{i \in R_{D3}} \omega_{1i}}{\sum_{i \in E_{D3}} \omega_{1i}} \qquad \text{(A1)}$$

Then the Wave III grand sample cross-sectional weight, before poststratification adjustment, is

$$\omega^{C}_{D3i} = \frac{\omega_{1i}}{r_{D3|1}} \qquad \text{(A2)}$$

For the poststratification adjustment, we divide an estimate of the total 1995 population in domain *D* who would be eligible in Wave III by an estimate of the same population from Wave III data as follows:

$$F^{C}_{D3} = \frac{\hat{X}_{E_{D3}}}{\sum_{i \in R_{D3}} \omega^{C}_{D3i}} \qquad \text{(A3)}$$

where

$$\hat{X}_{E_{D3}} = \sum_{i \in E_{D3}} \omega_{1i} \qquad \text{(A4)}$$

is the sum of the Wave I sample weight for the sample members who were eligible in Wave III (i.e., Wave I respondents who were alive at Wave III). Note that the estimate in Equation A4 may also be obtained by summing the Wave I sample weight for all sample members that have a Wave I sample weight, then subtracting the sum of the Wave I weight for all sample members that were deceased at Wave III. Tables 2 and 3 show the subpopulation estimates (A4) for the grade-sex-race poststratification domains.

Then the final Wave III grand sample weight, which is defined for all $I \in R_{D3}$ is

$$\varpi^{C}_{F,D3i} = F^{C}_{D3}\varpi^{C}_{D3i} \tag{A5}$$

The untrimmed cross-sectional weight was computed using Equation A1 to Equation A5. The trimmed cross-sectional weight was computed in the same manner, except that the nonresponse-adjusted weight resulting from Equation A2 was trimmed to a value of 6000 before being substituted into Equation A4 and Equation A5.

Grand Sample Longitudinal Weight

For each sample school, define the school-specific weighted response rate for Wave III given Wave II response as

$$r_{D3|2} = \frac{\sum_{i \in R_{D3}} \varpi_{1i}}{\sum_{i \in E_{D3}} \varpi_{1i}} \tag{A6}$$

Then the Wave III grand sample longitudinal weight, before poststratification adjustment, is

$$\omega^{L}_{D3i} = \frac{\omega_{2i}}{r_{D3|2}} \tag{A7}$$

which is defined for all $I \in R_{D2} \cap R_{D3}$.

For the poststratification adjustment, we divide an estimate of the total 1995 population in domain *D* who would be eligible for both Wave II and Wave III by an estimate of the same population from Wave III data as follows:

$$F^{L}_{D3} = \frac{\hat{X}_{E_{D2} \cap E_{D3}}}{\sum_{i \in R_{D2} \cap R_{D3}} \omega^{L}_{D3i}} \tag{A8}$$

where

$$\hat{X}_{E_{D2} \cap E_{D3}} = \sum_{i \in E_{D2} \cap E_{D3}} \omega_{2i} \tag{A9}$$

is the sum of the Wave II sample weight for the sample members who were eligible in Wave II and Wave III (i.e., Wave I and Wave II respondents who were alive at Wave III). Note that the estimate in Equation A9 may also be obtained by summing the Wave II sample weight for all sample members that have a Wave II sample weight, then subtracting the sum of the Wave II weight for all sample members that were deceased at Wave III. Tables 4 and 5 show the subpopulation estimates (A9) for the grade-sex-race poststratification domains.

Then the final Wave III grand sample longitudinal weight, which is defined for all $I \in R_{D2} \cap R_{D3}$, is

$$\omega^{L}_{F,D3i} = F^{L}_{D3}\omega^{L}_{D3i} \qquad \text{(A10)}$$

The untrimmed longitudinal weight was computed using Equation A6 to Equation A10. The trimmed longitudinal weight was computed in the same manner, except that the nonresponse-adjusted weight resulting from Equation A7 was trimmed to a value of 6000 before being substituted into Equation A9 and Equation A10.

Table 7 summarizes the total population estimates derived from all the sample weights.

Table 7. Weight Totals

Weights	Sample Members with Sample Weights		Sample Members Eligible (Alive) at Wave III	
	n	Sum of Weights	n	Sum of Weights
Wave I Grand Sample Weight	18,924	22,213,734	18,835*	22,106,340
Wave II Grand Sample Weight	13,568	18,824,889	13,516	18,754,449
Wave III Grand Sample Nonresponse-Adjusted Untrimmed Cross-Sectional Weight	14,322	22,106,340	14,322	22,106,340
Wave III Grand Sample Nonresponse-Adjusted Trimmed Cross-Sectional Weight	14,322	21,753,942	14,322	21,753,942
Wave III Grand Sample Poststratified Untrimmed Cross-Sectional Weight	14,322	22,104,088	14,322	22,104,088
Wave III Grand Sample Poststratified Trimmed Cross-Sectional Weight	14,322	22,104,088	14,322	22,104,088
Wave III Grand Sample Nonresponse-Adjusted Untrimmed Longitudinal Weight	10,828	18,754,449	10,828	18,754,449
Wave III Grand Sample Nonresponse-Adjusted Trimmed Longitudinal Weight	10,828	18,337,600	10,828	18,337,600
Wave III Grand Sample Poststratified Untrimmed Longitudinal Weight	10,828	18,754,448	10,828	18,754,448
Wave III Grand Sample Poststratified Trimmed Longitudinal Weight	10,828	18,754,448	10,828	18,754,448
Wave I Public-Use Sample Weight	6,504	22,261,000	6,476	22,144,325
Wave II Public-Use Sample Weight	4,834	18,817,312	4,814	18,730,298
Wave III Public-Use Sample Nonresponse-Adjusted Cross-Sectional Weight	4,882	22,144,325	4,882	22,144,325
Wave III Public-Use Sample Poststratified Cross-Sectional Weight	4,882	22,144,327	4,882	22,144,327
Wave III Public-Use Sample Nonresponse-Adjusted Longitudinal Weight	3,844	18,730,298	3,844	18,730,298
Wave III Public-Use Sample Poststratified Longitudinal Weight	3,844	18,730,296	3,844	18,730,296

* Of the 96 cases that were deceased at Wave III, seven cases did not have Wave I sample weights.

Appendix B: Informed Consent Forms

Interview

The National Longitudinal Study of Adolescent Health, or *Add Health,* is being conducted by the Research Triangle Institute for the Carolina Population Center at the University of North Carolina at Chapel Hill. The study is funded by the National Institutes of Health in Washington, DC.

Add Health is helping researchers understand the health of young adults and the behaviors that affect their health. In particular, it will promote understanding of how young adults' friendship networks and romantic relationships lead to choices that influence their health and well-being as they move into adulthood.

Because of your past participation in Add Health (or because of your relationship with an Add Health participant) you have been asked to take part in this interview. This means working with a professional interviewer to complete a questionnaire about your relationships and activities with others, such as friends, current or former spouses or cohabiting partners, dating partners, children, and parents. You will also be asked about your daily activities, schooling and employment, physical, emotional, and mental health, and health-related behaviors. In addition, the interviewer will measure your weight and height. These procedures will take about two hours.

You may skip any question you do not want to answer, you may decline to participate in any activity, and you may stop the interview at any time. Your answers will be held in strict privacy by the project staff and not given to unauthorized persons. Extensive security procedures are in place to make sure that participants' answers are not linked to their names. Your name will not appear in the interview file; these data are labeled only with an ID number. After the interview is finished, the interviewer will electronically transmit all ID information to the Add Health Security Manager in Toronto, Canada, and your interview answers to Research Triangle Institute. Storing different parts of the information in different places helps ensure that your identity will remain confidential.

If you have any questions about the Add Health study, you can call a toll-free number and speak with Dr. J. Richard Udry, the project director (1-877-377-9607). Add Health has been approved by the Institutional Review Board (IRB) on Research Involving Human Subjects at the University of North Carolina at Chapel Hill. The IRB can be reached at 1-919-966-3012. The project director believes there is minimal risk to you in being interviewed as part of this study.

You will receive $20 at the completion of the interview. Other than this reimbursement for your time, there is no specific benefit to you from completing an Add Health interview.

I have been given a copy of this consent form. I agree to participate in this Add Health interview.

_____ _____

Please Print Name Signature

_____

Date

Consent Form: Urine Specimen for STD Tests

You have been asked to provide a urine specimen that will be tested to see if you are infected with certain types of sexually transmitted diseases (STDs). Young adults are at very high risk for getting an STD but, since STDs often don't produce any symptoms, people may not know they are infected. Many STDs are easily cured by a few doses of an antibiotic. If an STD goes untreated, it can seriously damage one's health, particularly the ability to have children; harm unborn children; and be given to others. This is why it is important for you to be tested and to learn your results.

Your urine will be tested for gonorrhea and chlamydia using tests that have been approved as reliable by the Food and Drug Administration. The results of these tests will be made available to you. Your urine will also be stored and may be used later for other research purposes not limited to testing for STDs. Other test results, if any, will not be made available to you. Your urine will never be tested for drugs.

You are encouraged to call a toll-free number, 1-877-861-3633, beginning four weeks from today for the results of your chlamydia and gonorrhea tests. Although these tests are very good, they can be wrong. If you have a positive test result—that is, if the test indicates that you are infected with an STD—you will have an opportunity to talk with a trained counselor.

To provide the sample, you will urinate into a collection container, seal the container with the materials provided, and return it to the interviewer. The container will be labeled with an ID number that is different from the ID number associated with your interview responses. In this way, your name will not be linked to your specimen or test results.

If you provide a urine specimen, you will receive $10. The project director realizes that being tested for STDs may make you nervous or uncomfortable, but knowing your infection status is important. If you learn you are infected, you can seek simple medical treatment.

These procedures for urine collection and testing have been approved by the Institutional Review Board on Research Involving Human Subjects at the University of North Carolina at Chapel Hill (1-919-966-3012). If you have questions about these STD tests, you can call a toll-free number to speak with Dr. J. Richard Udry, the Add Health project director (1-877-377-9607). For general questions about STDs, you can call the free 24-hour Centers for Disease Control and Prevention National STD and AIDS Hotline (1-800-277-8922).

I have been given a copy of this consent form and an STD Information Sheet. I agree to provide a urine specimen and I consent to its testing and storage as described above.

_____ _____

Please Print Name Signature

_____

Date

Appendix C: Add Health Prisoner Interview Protocol

When in the course of locating Add Health respondents, it is determined that a respondent is incarcerated, the interviewer will contact by telephone the authorities responsible for controlling access at the place of incarceration, and determine the process for obtaining access to the respondent, including any documentation required by the authorities. We understand that, irrespective of UNC SPH IRB approval of our protocol, Add Health will be required to apply for IRB clearance on a case-by-case basis from the local IRB that governs each correctional facility housing a prospective Wave III respondent.

Immediately after a respondent has been located in a correctional facility, the interviewer will find out from prison administration when the respondent's incarceration is scheduled to end. If this date is within the period of Wave III fieldwork, the interviewer will seek permission from prison administration to privately ask the respondent for contact information after his or her release from prison. If the respondent provides this information, the field interviewer will contact the respondent after his or her release from prison to schedule the interview. If the prison administration requires special permission or local IRB clearance to access the respondent privately to get this information, Add Health will make every attempt to provide necessary documentation to obtain such permission.

If the incarceration will end after the Wave III data collection period has ended, Add Health will make every attempt to meet the requirements of the local IRB to gain access to the respondent in prison to conduct the interview, and will provide a description of the study protocol as well as this document and others indicating approval by our IRB. If local IRB clearance cannot be obtained, and the date of release is beyond the study data collection period, the respondent will not be interviewed.

After obtaining local IRB clearance, the interviewer will conduct the interview under conditions that provide him or her with safety and security, while at the same time precluding the prison staff from hearing or seeing the respondent's answers to questions. If the authorities are unwilling or unable to provide these conditions, the interview will not be conducted.

Interviewers will conduct face-to-face interviews with prisoners in their place of incarceration, using our IRB-approved laptop administration. In addition to the precautions that are in effect to protect the confidentiality of the information provided by respondents through the Add Health security management protocol, the interview will take place in a private place within the correctional facility. Prison administration will be required to agree that the privacy of the interview and its contents will not be invaded by the prison staff, or the interview will not be conducted.

The interviewer will describe the study protocol to the prisoner and will directly answer any questions about the study and the interview. The interviewer will carefully review with the prisoner informed consent forms for participation in the study and/or collection of biological specimens. The interviewer will make sure the prisoner understands that participation in the study is strictly voluntary, that all information provided will be held in strict confidentiality, and that participating in the study will in no way affect his or her chances for parole or treatment within the correctional facility. The interviewer will make certain that the informed consent forms have been read, understood, and signed before proceeding with the interview and/or biological specimen collection.

Prior to conducting the interview, Add Health field interviewers will determine whether prisoners are permitted to make unmonitored telephone calls from the prison facility. If prisoners are not permitted to make unmonitored telephone calls, respondents will not be asked to provide biological specimens for

diagnostic HIV and STD testing because they will not be able to confidentially access their test results from the Add Health toll-free number for results reporting.

Before collecting biological specimens, the Add Health field interviewers will determine whether adequate health services are available in the prison facility to retest and/or treat respondents who may test positive for either HIV or STDs. The Add Health informed consent process for the collection of biological samples will inform the respondents who provide biological specimens that they may be a risk of differential treatment by prison officials and/or fellow inmates if they seek health care services for HIV or STDs within the prison facility. The respondent will be able to refuse to provide biological specimens at any time during the interview process.

Before asking the respondent to provide a urine sample, the Add Health field interviewer will need to ascertain whether prison administration is willing to make available a restroom in close proximity to the place where the interview will take place. If the prison is not able to provide a restroom that will give the respondent privacy for providing a urine sample, the respondent will not be asked to provide a urine sample.

Through the Add Health informed consent process, prisoners will have understood that they can decline to answer any question or to participate in any part of the interview activities, and that they may end the interview at any time. Prisoners will also have understood that they may consent to being interviewed but decline to provide biological specimens.

Monetary incentives for completing the interview and providing biological specimens will be provided in accordance with prison requirements, probably in most cases by depositing a check or cash with prison administration to the credit of the prisoner, to be used by him or her under the conditions specified by the prison. Alternative arrangements, where permitted by law and the prison, may sometimes be used, for example, a direct bank deposit for the prisoner. Prisoners will be paid the same incentives paid to non-prisoners.

All questions related to suicide or suicide ideation will be omitted from the survey interview instrument that is administered to respondents who are in prison.

Prisoners will be omitted from the Add Health telephone follow-up survey. This omission does not deprive the prisoner of any benefit, and does not appreciably damage the study.

Appendix D: Event History Calendar Public Events

Year	Month	Public Event
1990	1	Communism falls in Yugoslavia
1990	2	Communism falls in USSR
1990	4	PRETTY WOMAN released
1990	5	Jim Henson dies
1990	7	Marion Barry smokes crack
1990	8	East/West Germany reunited
1990	11	Magic Johnson says he has AIDS
1990	12	Madonna's JUSTIFY MY LOVE
1991	1	Operation Desert Storm
1991	2	Cease fire ends Gulf War
1991	3	Eric Clapton's son dies
1991	4	Wm Kennedy Smith sex scandal
1991	5	TRUTH OR DARE released
1991	6	Mount Pinatubo erupts
1991	7	First Lallapalooza tour opens
1991	8	Lithuania/Latvia independent
1991	10	Anita Hill accuses Clarence Thomas
1991	11	Terry Waite freed in Lebanon
1991	12	Terrorists free Terry Anderson
1992	2	WAYNE'S WORLD released
1992	3	SILENCE OF THE LAMBS Oscar
1992	4	Rodney King verdict: LA riots
1992	5	Quayle blasts Murphy Brown
1992	6	Quayle misspells "potato"
1992	7	Barcelona Olympics open
1992	8	Dream Team wins gold
1992	9	Magic returns to the Lakers
1992	10	SEX by Madonna is published
1992	11	Bill Clinton elected president
1992	12	Charles and Di separate
1993	1	Clinton sworn in as president
1993	2	World Trade Center bombed
1993	4	Fire kills Branch Davidians
1993	5	Last episode of CHEERS
1993	6	Wife amputates Bobbitt's penis
1993	8	Letterman leaves NBC for CBS
1993	10	River Phoenix dies
1993	11	European Union formed
1993	12	NIRVANA UNPLUGGED airs on MTV

1994	1	Serbs pound Sarajevo
1994	2	Lillehammer Olympics begin
1994	3	SCHINDLER'S LIST wins Oscar
1994	4	Kurt Cobain commits suicide
1994	5	Paula Jones sues Bill Clinton
1994	6	NY Rangers win Stanley Cup
1994	7	FORREST GUMP released
1994	8	Woodstock 1994
1994	9	Baseball strike: no Series
1994	10	PULP FICTION opens in theaters
1994	11	Susan Smith says sons kidnpped
1994	12	Richard Gere/C. Crawford split
1995	1	Earthquake in Japan: 5,000 dead
1995	2	Transpacific balloon flight
1995	3	Americans to Mir Space Station
1995	4	Oklahoma City bombing
1995	5	Baseball returns after strike
1995	6	OJ's "gloves don't fit" ploy
1995	7	Heat wave kills 800
1995	8	Jerry Garcia dies
1995	9	Ripken breaks Gehrig's record
1995	10	Braves win World Series
1995	11	Beatles release "new" single
1995	12	Bosnia/Croatia sign treaty
1996	1	Dallas wins Super Bowl
1996	2	Suicide bomber in Sri Lanka
1996	3	Charles and Di divorce
1996	4	Ted Kaczynski is Unabomber
1996	5	Valujet crash in Everglades
1996	6	HUNCHBACK OF NOTRE DAME movie
1996	7	Atlanta's Olympic Park bombed
1996	8	Four women enter the Citadel
1996	9	Tupac Shakur fatally shot
1996	10	Yankees win World Series
1996	11	Bill Clinton re-elected
1996	12	SCREAM opens in theaters
1997	1	Ennis Cosby murdered
1997	2	STAR WARS re-released
1997	3	Hale-Bopp comet appears
1997	4	Tiger Woods wins Masters
1997	5	First Lilith Fair tour
1997	6	Bulls win 5th NBA title
1997	7	Pathfinder lands on Mars

1997	8	Princess Di dies in car crash
1997	9	Mother Teresa dies
1997	10	British nanny guilty of murder
1997	11	McCaughey septuplets born
1997	12	TITANIC opens in theaters
1998	1	Neb/Mich–football co-champions
1998	2	US jet cuts Italian ski cable
1998	3	Kentucky NCAA basketball champ
1998	4	Viagra on market
1998	5	Last SEINFELD episode
1998	6	Bulls win 6th NBA title
1998	7	France hosts/wins World Cup
1998	8	US embassies in Africa bombed
1998	9	McGwire breaks home-run record
1998	10	Matthew Shepard murdered
1998	11	Jesse Ventura elected governor
1998	12	Bill Clinton impeached
1999	1	Michael Jordan retires
1999	2	Jerry Falwell outs Teletubby
1999	3	Joe DiMaggio dies
1999	4	Columbine High School shooting
1999	5	STAR WARS: EPISODE 1 released
1999	6	TARZAN/SOUTH PARK in theaters
1999	7	JFK Jr/wife/sister plane crash
1999	8	Huge earthquake in Turkey
1999	9	Lauryn Hill MTV video award
1999	10	EgyptAir Flight 990 crashes
1999	11	Bonfire collapses at Texas A&M
1999	12	Puff Daddy/J. Lopez arrested
2000	1	World survives Y2K scare
2000	2	Charles Schultz dies
2000	3	AMERICAN BEAUTY Best Picture
2000	4	Elian reunited with father
2000	5	Love Bug computer virus
2000	6	Elian returns to Cuba with dad
2000	7	Concorde crashes near Paris
2000	8	Russian submarine Kursk sinks
2000	9	Abortion pill wins US approval
2000	10	Yankees win 26th World Series
2000	11	Presidential election
2000	12	Gore concedes to Bush
2001	1	Bush inaugurated as President