

National Longitudinal Study of Adolescent Health

Wave III Weights

DEVELOPED BY:

Paul P. Biemer, PhD
Elvessa D. Aragon-Logan, MS

RTI International
3040 Cornwallis Road
P.O. Box 12194
Research Triangle Park, NC 27709-2194

Add Health Wave III Weights

Introduction

In this document, 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:

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

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:

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

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.

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.

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 ineligible 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.

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 $[1 + CV(\text{weights})]$ where $CV22(\text{weights})$ 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.

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.

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
<i>Cases where person interviewed was not the correct Add Health respondent</i>								
1	494	Interview completed (all sections) <i>but</i> WAVE3 = !	1	!	!	13	E	NR
1	494	Interview completed (all sections) <i>but</i> WAVE3 = !	1	1	!	25	E	NR
<i>Other eligible nonrespondents</i>								
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 <i>but</i> 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

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

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

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

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

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

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

$$\omega_{D3i}^C = \frac{\omega_{1i}}{r_{D3|1}} \quad (A2)$$

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:

$$F_{D3}^C = \frac{\hat{X}_{E_{D3}}}{\sum_{i \in R_{D3}} \omega_{D3i}^C} \quad (A3)$$

where

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

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

$$\omega_{F,D3}^C = F_{D3}^C \omega_{D3}^C \quad (A5)$$

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

$$r_{D3|2} = \frac{\sum_{i \in R_{D3}} \omega_{1i}}{\sum_{i \in E_{D3}} \omega_{1i}} \quad (A6)$$

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

$$\omega_{D3i}^L = \frac{\omega_{2i}}{r_{D3|2}} \quad (A7)$$

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:

$$F_{D3}^L = \frac{\hat{X}_{E_{D2} \cap E_{D3}}}{\sum_{i \in R_{D2} \cap R_{D3}} \omega_{D3i}^L} \quad (A8)$$

where

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

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

$$\omega_{F,D3i}^L = F_{D3}^L \omega_{D3i}^L \quad (A10)$$

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.