

## Some Notes on AddHealth Sampling Procedures <sup>1</sup>

Following are notes on the sample design of the National Longitudinal Study of Adolescent Health. They were garnered from documents found at: <http://www.cpc.unc.edu/addhealth/>

### In-Home Sample Weights:

1. Cluster Variable – School ID at Wave I: **PSUSCID**

Freq	ID #
20372	001 to 371
373	999 (missing)

2. Strata Variable – Region **REGION**

Freq	Code	Region
4917	1	West
4956	2	Midwest
7445	3	South
2957	4	Northeast
470	!	Missing

### Design and Implementation of *In-School* Sample

- Initial sample 80 High Schools and 52 Feeder Schools = 132 total schools (nationally representative)
- High schools were sampled with probability proportional to school size
- Feeder schools – HS provided all schools expected to contribute at least 5 students to the entering class. A single feeder school was selected for each HS -- probability proportional to the % of HS's entering class that came from the feeder.
- 4 HSs had no feeder b/c they drew from so many; 20 HSs were their own FS b/c they contained 7<sup>th</sup> or 8<sup>th</sup> grade
- Thus, 56 FS selected, 4 refused = 52

### Design and Implementation of *Wave I In-Home* Sample

- Total N = 27559, made up of:
  - o Core = 16,044 students enrolled in G7-12
  - o Saturated samples from 2 HS = 3350 (PAIRS sample schools)
  - o Two groups of supplemental samples:
    - a. Non-genetic – various ethnic groups and disabled
    - b. Genetic – sibling relationships

### Core Sample Selection

- Roughly equal-sized samples drawn from 12 student-level strata (formed by cross-classifying students by their sex and grade)
- Stratum sample size targets = 16,044 (total core) / 12 = 1333 selections per stratum
- School-level targets set for each stratum = 1333 / # schools with at least one student in the stratum. E.g. if 75 schools included 9<sup>th</sup> grade males, the target sample size for 9<sup>th</sup> grade males at each of these schools = 18 (1333/75).
- Missing grade or sex info had to be imputed in some cases to allow classification of each student by stratum.
- In the end, some variation of sample sizes in strata occurred (range: 1223 to 1398)
- All of the students at the two PAIRS schools were selected for In-home data collections. In those schools, the core sample cases were a subsample of the PAIRS sample.

---

<sup>1</sup>Prepared by Liz Mogford. November 29, 2000. C:\all\help\helpnew\AHSamp1.wpd