

**Report prepared by**

Bria Gresham

Theresa Osypuk

Waves I-V Neighborhood Concentrated Disadvantage Data for Longitudinal Analyses User Guide



CAROLINA POPULATION CENTER | CAROLINA SQUARE - SUITE 210 | 123 WEST FRANKLIN STREET | CHAPEL HILL, NC 27516

Add Health is supported by cooperative agreement U01 AG071448 (Hummer) and cooperative agreement U01 AG071450 (Hummer and Aiello) from the National Institute on Aging at the National Institutes of Health (NIH), with co-funding provided NICHD, NIMHD, NIDA, OBSSR, and ODP at NIH.

Acknowledgments

Add Health is funded by grant P01 HD31921 (Harris) from the Eunice Kennedy Shriver National Institute of Child Health and Human Development (NICHD), with cooperative funding from 23 other federal agencies and foundations. Add Health is currently directed by Robert A. Hummer and funded by the National Institute on Aging cooperative agreements U01 AG071448 (Hummer) and U01AG071450 (Aiello and Hummer) at the University of North Carolina at Chapel Hill. Add Health was designed by J. Richard Udry, Peter S. Bearman, and Kathleen Mullan Harris at the University of North Carolina at Chapel Hill.

The contextual data linkage was supported by grants to Bria Gresham from the Minnesota Population Center, the Institute of Child Development, and the Gunnar Laboratory for Developmental Psychobiology Research at the University of Minnesota.

The authors gratefully acknowledge support from Lindsey Fabian and Dr. David van Riper for their guidance with this ancillary study and the reviewers of this Add Health ancillary study for their suggestions.

Citation for User Guide

Gresham, B., and T. Osypuk. 2025. ***Waves I-V Neighborhood Concentrated Disadvantage Data for Longitudinal Analyses User Guide***. Chapel Hill, NC: Carolina Population Center, University of North Carolina at Chapel Hill. Available from: <https://doi.org/10.17615/gvf4-8e30>

Introduction

This document provides an overview of the Add Health ancillary study entitled, “Wave I-V Neighborhood Concentrated Disadvantage Data for Longitudinal Analyses” and its accompanying dataset. Five indicators of neighborhood concentrated disadvantage are linked to Add Health participants at the block group level to reflect concentrated disadvantage.

1. Proportion of persons below the federal poverty level (last year)
2. Proportion of persons 16+ years old in the labor force and unemployed
3. Proportion of persons 25+ years old with less than a high school diploma
4. Proportion of households on public assistance (last year)
5. Proportion of households with children that are single-headed

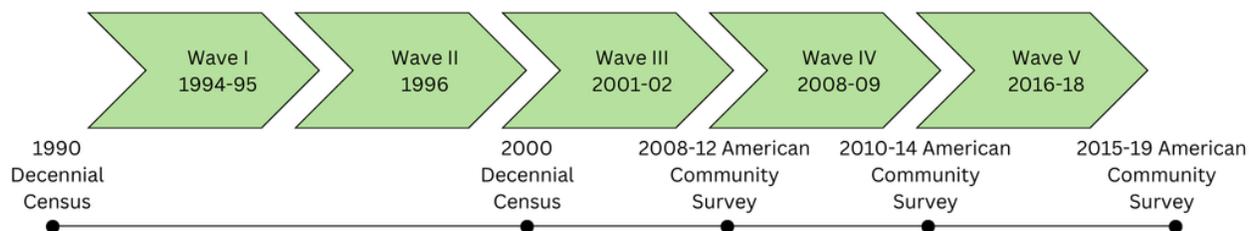
These data differ from existing contextual data available from Add Health in several key ways, such that 1) the census block group boundaries are standardized to 2010 boundaries, 2) the values of each indicator variable are linearly interpolated to estimate the values of intercensal years and align with the year of Add Health assessment (e.g., 1994 values of the proportion of persons below the federal poverty line were estimated from 1990 and 2000 decennial census values), and 3) weighted composite variables of neighborhood concentrated disadvantage are included. To calculate the composite neighborhood concentrated disadvantage variables, a series of principal component analyses were conducted to extract weights for each indicator. Each indicator was then multiplied by the extracted weight, and the weighted items were then summed.

Data Structure and Form

The data file (**w5nbrhddisad**) contains a total of 31 variables. The first variable is the respondent identifier (AID), by which these data can be merged with other Add Health data files. The remaining variables include the indicator and weighted composite score variables for each Wave I through V. With the exception of the respondent identifier, variable values ranged from 0-1 (see below for codes used to indicate missingness).

Source Description

The raw data in the dataset were obtained from IPUMS National Historical Geographic Information System (NHGIS). Relevant data included data tables for the five indicator variables of interest from the 1990 decennial census, 2000 decennial census, 2008-12 American Community Survey, 2010-14 American Community Survey, and 2015-19 American Community Survey, as these data represented the census data most proximal the timing of the Add Health assessments. In addition, 1990 and 2000 census block group/block group part crosswalks were obtained from IMPUS NHGIS. The figure below shows the data sources used to calculate the linearly interpolated values, with the two most proximal data sources used for each Add Health time point.



Variable Naming Conventions

All of the variables in the dataset adhere to Add Health naming conventions. Each of the contextual variables in the dataset begin with the letter “B”, representing the geographic level, for census block group based on the respondents’ residential address. The second character reflects the wave number (i.e., 1, 2, 3, 4, or 5). Characters 3+ describe the construct reflected by the variable.

Data Dictionary

The name and a brief description of each variable in the dataset are presented below.

Variable Name	Variable Description
AID	Add Health respondent identifier
B1PRPPOVERTY	Proportion of persons below the federal poverty level last year
B1PRPUNEMP	Proportion of persons 16+ years old in the labor force and unemployed
B1PRPNODIPL	Proportion of persons 25+ years old with less than high school diploma
B1PRPPUBASST	Proportion of households on public assistance last year
B1PRPSNGHEAD	Proportion of households with children that are single headed
B1NBRHDDISAD	Neighborhood concentrated disadvantage
B2PRPPOVERTY	Proportion of persons below the federal poverty level last year
B2PRPUNEMP	Proportion of persons 16+ years old in the labor force and unemployed
B2PRPNODIPL	Proportion of persons 25+ years old with less than high school diploma
B2PRPPUBASST	Proportion of households on public assistance last year
B2PRPSNGHEAD	Proportion of households with children that are single headed
B2NBRHDDISAD	Neighborhood concentrated disadvantage
B3PRPPOVERTY	Proportion of persons below the federal poverty level last year
B3PRPUNEMP	Proportion of persons 16+ years old in the labor force and unemployed
B3PRPNODIPL	Proportion of persons 25+ years old with less than high school diploma
B3PRPPUBASST	Proportion of households on public assistance last year
B3PRPSNGHEAD	Proportion of households with children that are single headed
B3NBRHDDISAD	Neighborhood concentrated disadvantage
B4PRPPOVERTY	Proportion of persons below the federal poverty level last year
B4PRPUNEMP	Proportion of persons 16+ years old in the labor force and unemployed
B4PRPNODIPL	Proportion of persons 25+ years old with less than high school diploma
B4PRPPUBASST	Proportion of households on public assistance last year
B4PRPSNGHEAD	Proportion of households with children that are single headed
B4NBRHDDISAD	Neighborhood concentrated disadvantage
B5PRPPOVERTY	Proportion of persons below the federal poverty level last year
B5PRPUNEMP	Proportion of persons 16+ years old in the labor force and unemployed
B5PRPNODIPL	Proportion of persons 25+ years old with less than high school diploma
B5PRPPUBASST	Proportion of households on public assistance last year
B5PRPSNGHEAD	Proportion of households with children that are single headed
B5NBRHDDISAD	Neighborhood concentrated disadvantage

Missing codes

This dataset has three different codes that indicate the source of missing data.

- -9992 Missing in the source data
- -9991 Respondent was not interviewed in that wave
- -9990 Respondent lacks the geocode necessary for merging the source data

Notes

Data from the 2008-12, 2010-14, and 2015-19 American Community Surveys were not geographically standardized prior to interpolation, as these data already represented 2010 census boundaries. As the American Community Survey represents 5-year averages, data from these sources were anchored to the mid-point of the 5-year interval prior to interpolation (e.g., 2010 was used as the representative year for 2008-12 American Community Survey).