Add Health
The National Longitudinal Study of Adolescent to Adult Health

Wave IV County Health and Mozbility Data Documentation

Report prepared by

Taylor W. Hargrove

Lauren Gaydosh

Max Reason

Daniel W. Belsky

UNC | CAROLINA POPULATION CENTER

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

Suggested Citation

Hargrove, Taylor W., Lauren Gaydosh, Max Reason, Daniel W. Belsky, 2019. The National Longitudinal Study of Adolescent to Adult Health (Add Health), Wave IV County Health and Mobility Data Documentation. Chapel Hill, NC: Carolina Population Center, University of North Carolina at Chapel Hill. https://doi.org/10.17615/znj8-nc52

Acknowledgments

Introduction

The Wave IV County Health and Mobility database summarizes the socioeconomic, health, and mobility characteristics of the environments in which Add Health participants were living at the time of their Wave IV interview. County-level data describe (1) levels of and trends in chronic disease (hypertension, type-2 diabetes) and health risk behaviors (obesity, smoking, alcohol use); and (2) economic opportunity and inequality. This contextual database permits innovative research that investigates how place influences health, behavior, and social outcomes during the transition from adolescence to adulthood, thereby, enhancing studies of the determinants and sequelae of socio-geographic mobility. Comprehensively, the database provides indicators of contextual conduciveness to the socio-geographic and health-geographic mobility of Add Health participants from Wave I to Waves IV and V.

Data

The following is a list of data that were collected from secondary data sources and merged to Wave IV of Add Health. These variables are available at the county or state level. Data were matched to the county or state that the Add Health respondent was living in at the time of the Wave IV interview and data were matched to respondents so as to insure that these contextual variables correspond as closely as possible to the year in which the Add Health respondents were interviewed at Wave IV (2008).

Life Expectancy and Mortality Risk

Data on life expectancy and mortality risk come from the Institute for Health Metrics and Evaluation (IHME). Annual county-level life tables were constructed using small area estimation methods from de-

identified death records from the National Center for Health Statistics (NCHS), and population counts from the US Census Bureau, NCHS, and the Human Mortality Database.

County level life expectancy (2005) C4LE001
County level mortality risk age 0-5 (2005) C4LE002
County level mortality risk age 5-25 (2005) C4LE003
County level mortality risk age 25-45 (2005) C4LE004
County level mortality risk age 45-65 (2005) C4LE005
County level mortality risk age 65-85 (2005) C4LE006
% change in life expectancy, 1980-2014 .... C4LE007

Citation: Lindgren et al. 2017 *JAMA Internal Medicine* [Inequalities in Life Expectancy Among US Counties, 1980-2014](#) [1]

Diabetes Prevalence

Data on diabetes prevalence came from the IHME. IHME researchers used a two-stage modeling procedure. In the first stage, self-reported and biomarker data from National Health and Nutrition Examination Survey (NHANES) were used to build models for predicting true diabetes status, which were applied to impute true diabetes status for respondents in the BRFSS. In the second stage, small area models were fit to imputed Behavioral Risk Factor Surveillance System (BRFSS) data to derive county-level estimates of diagnosed, undiagnosed, and total diabetes prevalence, as well as rates of diabetes diagnosis and effective treatment.

County level age-standardized diagnosed diabetes prevalence: Total (2008) C4DI001T
County level age-standardized diagnosed diabetes prevalence: Male (2008) C4DI001M
County level age-standardized diagnosed diabetes prevalence: Female (2008) C4DI001F
County level age-standardized undiagnosed diabetes prevalence: Total (2008) C4DI002T
County level age-standardized undiagnosed diabetes prevalence: Male (2008) C4DI002M
County level age-standardized undiagnosed diabetes prevalence: Female (2008) C4DI002F
County level age-standardized total diabetes prevalence: Total (2008) C4DI003T
County level age-standardized total diabetes prevalence: Male (2008) C4DI003M
County level age-standardized total diabetes prevalence: Female (2008) C4DI003F
County level age-standardized diabetes awareness: Total (2008) C4DI004T
County level age-standardized diabetes awareness: Male (2008) C4DI004M
County level age-standardized diabetes awareness: Female (2008) C4DI004F
County level age-standardized diabetes control: Total (2008) C4DI005T
County level age-standardized diabetes control: Male (2008) C4DI005M
County level age-standardized diabetes control: Female (2008) C4DI005F
% change 1999-2012 County level age-standardized diagnosed diabetes prevalence: Total C4DI006T
% change 1999-2012 County level age-standardized diagnosed diabetes prevalence: Male C4DI006M
% change 1999-2012 County level age-standardized diagnosed diabetes prevalence: Female C4DI006F
% change 1999-2012 County level age-standardized undiagnosed diabetes prevalence: Total C4DI007T
% change 1999-2012 County level age-standardized undiagnosed diabetes prevalence: Male .. C4DI007M
% change 1999-2012 County level age-standardized undiagnosed diabetes prevalence: Female C4DI007F
% change 1999-2012 County level age-standardized total diabetes prevalence: Total C4DI008T
% change 1999-2012 County level age-standardized total diabetes prevalence: Male C4DI008M
% change 1999-2012 County level age-standardized total diabetes prevalence: Female C4DI008F
% change 1999-2012 County level age-standardized diabetes awareness: Total C4DI009T

% change 1999-2012 County level age-standardized diabetes awareness: Male C4DI009M
% change 1999-2012 County level age-standardized diabetes awareness: Female C4DI009F
% change 1999-2012 County level age-standardized diabetes control: Total C4DI0010T
% change 1999-2012 County level age-standardized diabetes control: Male C4DI0010M
% change 1999-2012 County level age-standardized diabetes control: Female C4DI0010F

Citation: Dwyer-Lindgren et al. 2016 *Diabetes Care* [Diagnosed & Undiagnosed Diabetes Prevalence by County in the US 1999-2012](#) [2]

Drinking Patterns

Data on drinking patterns came from the IHME. IHME researchers applied small area models to BRFSS data on self-reported drinking, incorporating spatial and temporal smoothing.

County level age-standardized prevalence of any drinking: Total (2008) C4DR001T
County level age-standardized prevalence of any drinking: Male (2008) C4DR001M
County level age-standardized prevalence of any drinking: Female (2008)...............................C4DR001F
County level age-standardized prevalence of heavy drinking: Total (2008) C4DR002T
County level age-standardized prevalence of heavy drinking: Male (2008) C4DR002M
County level age-standardized prevalence of heavy drinking: Female (2008)..............................C4DR002F
County level age-standardized prevalence of binge drinking: Total (2008) C4DR003T
County level age-standardized prevalence of binge drinking: Male (2008) C4DR003M
County level age-standardized prevalence of binge drinking: Female (2008) C4DR003F
% change 2002-2012 County level age-standardized prevalence of any drinking: Total C4DR004T
% change 2002-2012 County level age-standardized prevalence of any drinking: Male C4DR004M
% change 2002-2012 County level age-standardized prevalence of any drinking: Female C4DR004F
% change 2002-2012 County level age-standardized prevalence of heavy drinking: Total C4DR005T
% change 2002-2012 County level age-standardized prevalence of heavy drinking: Male C4DR005M
% change 2002-2012 County level age-standardized prevalence of heavy drinking: Female C4DR005F
% change 2002-2012 County level age-standardized prevalence of binge drinking: Total C4DR006T
% change 2002-2012 County level age-standardized prevalence of binge drinking: Male C4DR006M
% change 2002-2012 County level age-standardized prevalence of binge drinking: Female C4DR006F

Citation: Dwyer-Lindgren et al. 2015 *AJPH* [Drinking Patterns in US Counties from 2002-2012](#) [3]

Smoking Patterns

Data on smoking patterns came from the IHME. IHME researchers used data on 4.7 million adults age 18 and older from the BRFSS from 1996 to 2012. They derived cigarette smoking status from self-reported data in the BRFSS and applied validated small area estimation methods to generate estimates of current total cigarette smoking prevalence and current daily cigarette smoking prevalence.

County level prevalence of people who currently smoke: Total (2008).......................................C4SM001T
County level prevalence of people who currently smoke: Male (2008)..................................... C4SM001M
County level prevalence of people who currently smoke: Female (2008) C4SM001F
County level prevalence of people who currently smoke daily: Total (2008) C4SM002T
County level prevalence of people who currently smoke daily: Male (2008) C4SM002M
County level prevalence of people who currently smoke daily: Female (2008)...........................C4SM002F

Mean annualized rate of change from 1996-2012 in County level prevalence of people who currently smoke: Total C4SM003T
Mean annualized rate of change from 1996-2012 in County level prevalence of people who currently smoke: Male C4SM003M
Mean annualized rate of change from 1996-2012 in County level prevalence of people who currently smoke: Female C4SM003F
Mean annualized rate of change from 1996-2012 in County level prevalence of people who currently smoke daily: Total C4SM004T
Mean annualized rate of change from 1996-2012 in County level prevalence of people who currently smoke daily: Male C4SM004M
Mean annualized rate of change from 1996-2012 in County level prevalence of people who currently smoke daily: Female C4SM004F

Citation: Dwyer-Lindgren et al. 2014 *Population Health Metrics* Cigarette Smoking Prevalence in US Counties 1996-2012 [4]

Physical Activity and Obesity

Data on physical activity and obesity came from the IHME. Body mass index (BMI) is calculated from self-reported weight and height in BRFSS, adjusting for self-reporting bias using NHANES. Physical activity—both any physical activity and physical activity meeting recommended levels—is calculated from self-reported data in BRFSS. To generate estimates of obesity and physical activity prevalence for each county, IHME researchers used validated small area estimation methods.

County level prevalence of obesity: Male (2008) C4OP001M
County level prevalence of obesity: Female (2008).. C4OP001F
County level prevalence of physical activity: Male (2008) C4OP002M
County level prevalence of physical activity: Female (2008).. C4OP002F

Citation: Dwyer-Lindgren et al. 2013 *Population Health Metrics* Prevalence of physical activity and obesity in US counties, 2001–2011: a road map for action [5]

Hypertension Prevalence

Data on hypertension prevalence came from the IHME. Combing data from NHANES and BRFSS among adults age 30 and older, hypertension was defined as systolic blood pressure (BP) of at least 140 mm Hg, self-reported use of antihypertensive treatment, or both. Hypertension control was defined as systolic BP less than 140 mm Hg.

County level total hypertension prevalence: Male (2009) C4HT001M
County level total hypertension prevalence: Female (2009)..C4HT001F
County level self-reported hypertension prevalence: Male (2009)... C4HT002M
County level self-reported hypertension prevalence: Female (2009) C4HT002F
County level treated hypertension prevalence: Male (2009).. C4HT003M
County level treated hypertension prevalence: Female (2009) C4HT003F
County level controlled hypertension prevalence: Male (2009) C4HT004M
County level controlled hypertension prevalence: Female (2009)...C4HT004F
County level uncontrolled hypertension prevalence: Male (2009) C4HT005M
County level uncontrolled hypertension prevalence: Female (2009) C4HT005F
County level hypertension awareness prevalence: Male (2009)... C4HT006M

County level hypertension awareness prevalence: Female (2009) C4HT006F

Citation: Olives et al. 2013 *PLOS One* [Prevalence, awareness, treatment, and control of hypertension in United States counties, 2001-2009](#) [6]

County Health Ranking Quartiles

Data on county health ranking quartiles came from the Robert Wood Johnson Foundation's County Health Rankings and Roadmaps Project. Data characterize US counties in terms of overall health and range of health problems and health behaviors during 2010-2017.

County health outcome ranking quartile – compiles length of life and quality of life data from the NCHS, and reports of poor or fair health, poor physical health days, and poor mental health days from the BRFSS.

County health factor ranking quartile - includes health behaviors (smoking, diet, exercise, alcohol and drug use, and sexual activity), clinical care (access to care, quality of care), social and economic factors (education, employment, income, family and social support, and community safety), and physical environment (air and water quality, housing and transit).

County health outcome ranking quartile (2010) C4HR001
County health factor ranking quartile (2010) C4HR002

Citation: [Robert Wood Johnson Foundation's County Health Rankings and Roadmaps Project](#)

Measures of Mobility

Data on measures of mobility come from the Equality of Opportunity Project. These data characterize US counties in terms of resident socioeconomic mobility. The Equality of Opportunity Project draws on federal income tax data to create measures of mobility.

County relative mobility –slope from OLS regression of child rank on parent rank within each county in core sample using baseline income definitions. Correlation of the percentile rank in the national income distribution for children (based on average incomes between 2010 and 2012 for the 1980-1982 birth cohort) and their parents (whose income was measured over 1996-2000).

County absolute mobility – expected rank of children whose parents are at the 25th percentile of the national income distribution based on the rank-rank regression

Causal effect of county of childhood residence on adult household income –measured as the percentage gain or loss in income at age 26 caused by spending one additional year of childhood in a given county relative to the national mean for children born to a family earning an income of approximately $30,000 (the 25th percentile of the income distribution). Children were assigned to a county based on their location at age 16 (no matter where they live as adults), so that their location represents where they grew up.

County Gini coefficient – a measure of the amount of parental income inequality within commuting zones in the US. The following equation is used to compute the Gini coefficient:

$$Gini = \frac{2}{X_c} Cov(X_{ic}, P_{ic}),$$

where $\overline{X_c}$ is the mean family income (for 1996-2000) of parents in CZ c, and $Cov(X_{ic}, P_{ic})$ is the covariance between income level (X_{ic}) and percentile rank (P_{ic}) of parents in CZ c.

County relative mobility..C4EC001
County absolute mobility..C4EC002
Causal effect of county of childhood residence on adult household income at p25: Total C4EC003T
Causal effect of county of childhood residence on adult household income at p25: Male C4EC003M
Causal effect of county of childhood residence on adult household income at p25: Female C4EC003F
County Gini coefficient C4EC004

Citations: Chetty & Hendren 2018 The Effects of Neighborhoods on Intergenerational Mobility I: Childhood Exposure Effects [7]

Chetty & Hendren 2018 The Effects of Neighborhoods on Intergenerational Mobility II: County-level Estimates [8]
Chetty, Hendren, Kline & Saez 2014 Where is the Land of Opportunity? The Geography of Intergenerational Mobility in the United States [9]

Health Insurance Coverage

Data on health insurance coverage came from the Kaiser Family Foundation State Health Facts dataset. Information on the proportion of a population within a state that is covered by different types of health insurance coverage is available in this dataset across multiple years.

It is important to note that the Kaiser Family Foundation variables, which measure State-level insurance coverage, are reported at respondent Wave IV spatial locations despite being measured in years 2014, 2015, and 2016. Most Wave IV interviews occurred in 2008, with a total of 2% being administered in 2007 and 2009. Any analysis considering the inclusion of these variables needs to assess how accurately they represent respondent context at Wave IV. This involves the following measures:

Medicaid: Includes those covered by Medicaid, the Children's Health Insurance Program (CHIP), and those who have both Medicaid and another type of coverage, such as dual eligibles who are also covered by Medicare. A person having Medicaid coverage in the first half of the year but employer-based coverage in the last months of the year would be categorized as having Medicaid coverage in this analysis.

Medicare: Includes those covered by Medicare, Medicare Advantage, and those who have Medicare and another type of non-Medicaid coverage where Medicare is the primary payer. Excludes those with Medicare Part A coverage only and those covered by Medicare and Medicaid (dual eligibles).

Employer: Includes those covered by employer-sponsored coverage either through their own job or as a dependent in the same household.

Other Public: Includes those covered under the military or Veterans Administration.

Non-Group: Includes individuals and families that purchased or are covered as a dependent by non-group insurance.

Uninsured: Includes those without health insurance and those who have coverage under the Indian Health Service only.

State level proportion of total population covered by Medicaid health insurance: 2014 ....S4HI0114
State level proportion of total population covered by Medicaid health insurance: 2015 ....S4HI0115
State level proportion of total population covered by Medicaid health insurance: 2016 ....S4HI0116
State level proportion of total population covered by Medicare health insurance: 2014 ....S4HI0214
State level proportion of total population covered by Medicare health insurance: 2015 ....S4HI0215
State level proportion of total population covered by Medicare health insurance: 2016 ....S4HI0216
State level proportion of total population covered by Employer health insurance: 2014 ....S4HI0314
State level proportion of total population covered by Employer health insurance: 2015 ....S4HI0315
State level proportion of total population covered by Employer health insurance: 2016 ....S4HI0316
State level proportion of total population covered by non-group health insurance: 2014 ....S4HI0414
State level proportion of total population covered by non-group health insurance: 2015 ....S4HI0415
State level proportion of total population covered by non-group health insurance: 2016 ....S4HI0416
State level proportion of total population covered by other public health insurance: 2014 ....S4HI0514
State level proportion of total population covered by other public health insurance: 2015 ....S4HI0515
State level proportion of total population covered by other public health insurance: 2016 ....S4HI0516
State level proportion of total population uninsured: 2014 ....S4HI0614
State level proportion of total population uninsured: 2015 ....S4HI0615
State level proportion of total population uninsured: 2016 ....S4HI0616

Citation: Kaiser Family Foundation State Health Facts

Tax Burden on Tobacco

Data come from the Tax Burden on Tobacco dataset.

State level cigarette tax per pack, in dollars (2008) S4CT001

Citation: Citation: Orzechowski and Walker, 2016 Tax Burden on Tobacco [10]

Missing codes

The final digit of the missing codes indicates the reason for which they are missing. Missing codes that end in 2 (Ex. 92, 992) denote that information for that variable was not available in the source dataset. Missing codes that end in 8 (Ex. 98, 998) denote respondents in Add Health who lack the geocodes necessary for merging respondent locations to the various source data.