Producing Synthetic Estimates of Children's Health and Well-Being for Local Areas

Mark Mather, Beth Jarosz, and Linda A. Jacobsen

Population Reference Bureau

Paper prepared for the 2015 Population Association of America Annual Meeting
May 2015

Acknowledgement: This research was conducted with funding from and in partnership with the Child and Adolescent Health Measurement Initiative, based in the Bloomberg School of Public Health at Johns Hopkins University in Baltimore, Maryland.

BACKGROUND

Public health agencies need good data to make informed policy decisions that can improve the well-being of children. However, while child health data are readily available at the national level, city- and county-level health departments face numerous constraints in obtaining generalizable, valid, and reliable child health data that are needed to conduct community-based needs assessments: The sample sizes of most national surveys do not permit analyses of data at the state or local levels; local geographic identifiers are often restricted to protect respondent confidentiality; and budget constraints limit federal and state funds available to conduct surveys in local communities.

The National Survey of Children's Health (NSCH) includes over one hundred indicators of children's physical, mental and behavioral health status, access to quality health care and a medical home, as well as family, neighborhood, and social contextual factors. These data provide a rich resource for child and adolescent programmatic and policy decisions. However, sub-state data are unavailable in the NSCH and the number of respondents in the survey is insufficient to produce reliable local area estimates.

In this paper, we summarize the methods and results of a project that combines data from the NSCH and the Census Bureau's American Community Survey (ACS) to produce synthetic local area estimates of children's health and well-being. The synthetic local area estimates are constructed by applying state prevalence rates, broken down by race/ethnicity and income, to race- and income-specific population data at the substate level. We discuss several methods used to evaluate the accuracy and biases associated with these estimates.

DATA

Prevalence rates for this analysis are derived from the 2011-2012 NSCH, which is a national, representative survey that provides a broad range of information about children's health and well-being collected in a manner that allows comparisons among states. Telephone numbers are called at random to identify households with one or more children under 18 years old. In each household, one child was randomly selected to be the subject of the interview. A total of 95,677 surveys were completed nationally for children under 18 years old, and survey results are weighted to represent the population of non-institutionalized children ages 0-17 nationally and in each state. The purpose of the survey is to estimate national and state-level prevalence for a variety of physical, emotional, and behavioral child health indicators in combination with information on the child's family context and neighborhood environment.

Local population data are from the 2010-2012 three-year American Community Survey (ACS) microdata file. The ACS is a nationwide, continuous survey designed to provide communities with reliable and timely demographic, housing, social, and economic data every year. As of 2012, the ACS samples 3.5 million addresses each year, resulting in approximately 2.4 million final interviews.[1] The ACS combines population data across multiple years to produce reliable estimates for small counties, neighborhoods, and other local areas. While the ACS collects information on a wide range of social and economic issues, health data are limited to a single question about health insurance coverage and six questions about disabilities. By combining the local demographic and economic data from the ACS with the rich health data from the NSCH, we can create a unique profile of children's health outcomes for local areas.

The Census Bureau's American FactFinder provides a wealth of pretabulated information about the U.S. child population in local areas. However, a drawback of using published ACS data is that it's not possible to customize income and racial/ethnic categories to match the standard categories available from the NSCH. Data in American FactFinder are currently available for children living above or below the official poverty threshold for eight racial/ethnic groups:

- White, Non-Hispanic alone;
- Black/African American alone;
- American Indian/Alaska Native alone;
- Asian alone;
- Native Hawaiian/Other Pacific Islander alone;
- Some Other Race Alone;
- Two or More Races; and
- Hispanic/Latino.

In contrast, published NSCH data are readily available for four family income groups and four racial groups cross tabulated by Hispanic origin. Table 1 shows the estimated prevalence rates for overweight or obese children ages 10 to 17 in Maryland, by race/ethnicity and family income as a percentage of the Federal Poverty Level (FPL).[2] In some cases, we had to substitute a regional prevalence rate for the state rate because of large sampling errors associated with the NSCH data for Maryland. Childhood obesity varied widely, from 18 percent among higher-income white children to more than 50 percent among lower-income African American children.

Table 1: Overweight/Obese Rates for Children in Maryland (from the NSCH)

NSCH Prevalence Rate MARYLAND		0-99% FPL	100-199% FPL	200-399% FPL	400% FPL or Higher
Hispanic		50.0%*	45.7%*	36.2%*	23.8%*
White, non-Hispanic		40.9%*	33.6%*	24.1%	18.0%
Black, non-Hispanic		52.7%*	80.2%	36.3%	34.7%
Other, non-Hispanic		37.9%*	35.1%*	35.9%*	20.9%*

*Regional (South) rate substituted for state rate.
Source: 2011-2012 National Survey of Children's Health.

These results show that within racial/ethnic groups, there are wide differences in obesity rates across different family income categories. Children in the highest income group have lower obesity rates than those in lower-income categories, regardless of racial/ethnic identification. Similarly, within different income groups, obesity rates vary across racial/ethnic groups, with lower rates among white children and higher rates among African American children. Given these patterns, applying NSCH prevalence rates that differentiate children by both race/ethnicity and family income should yield more accurate synthetic estimates than rates based on race/ethnicity or income alone.

Table 2 presents ACS data on the population under age 18 living in Baltimore City, Maryland cross tabulated by race/ethnicity and income. The majority of children living in Baltimore are African American—two thirds of whom live in families with income below 200% of the Federal poverty level.

Table 2: Child Population Ages 10 to 17 in Baltimore (from the ACS)

ACS Child Pop. BALTIMORE	TOTAL	0-99% FPL	100-199% FPL	200-399% FPL	400% FPL or Higher
TOTAL	**54,028**	19,094	14,000	13,959	6,975
Hispanic	2,003	777	463	548	215
White, non-Hispanic	7,838	1,300	1,307	2,348	2,883
Black, non-Hispanic	41,534	16,310	11,532	10,335	3,357
Other, non-Hispanic	2,653	707	698	728	520

Source: 2010-2012 American Community Survey.

We acquired population data that align with NSCH racial/ethnic and family income groups by using the Census Bureau's ACS microdata files. The microdata files allowed us to produce custom estimates and the associated standard errors for indicators and geographic areas that are unavailable through the Census Bureau's published tables. Three-year estimates from the ACS are available for geographic areas with at least 20,000 people. However, in order to increase the reliability of the data, we limited our analysis to the 583 counties and 297 cities with populations of at least 100,000 in 2012. Our population estimates are slightly lower than the Census Bureau's published estimates of the population under age 18 because we limited our analysis to children for whom poverty status is determined.

We chose to use the three-year ACS data, rather than the five-year ACS data, because the years (2010-2012) were better aligned with those from the NSCH (2011-2012). However, the five-year data from 2009-2013 could potentially be used to produce estimates for even smaller geographic areas.

METHODS

Synthetic estimation is a procedure that can be used to combine demographic data for local geographic areas with prevalence rates from a "parent" geography to create estimated prevalence rates for local areas. In effect, synthetic estimation involves reweighting "parent" prevalence rates based on the unique demographic and income characteristics of a local area. In our analysis, the weights are equal to the share of the total population in each of the 16 combined race/ethnicity and family income categories shown in Table 2.

Synthetic estimation was first introduced in the United States by the National Center for Health Statistics, which used a synthetic estimation procedure to produce state-level disability estimates.[3] More recently, the Child and Adolescent Health Measurement Initiative described how synthetic estimates could be used to construct local estimates of child well-being based on NSCH data.[4] This is the first project that applies NSCH prevalence rates to local population estimates on a broad scale.

Synthetic estimates have been widely used because they are relatively simple to produce, inexpensive, and easy to explain to nontechnical audiences. However, from a statistical standpoint, synthetic estimates are inherently biased because the derived estimates are expected to differ to some degree from the true value of the parameter being measured. Our assumption in creating synthetic estimates for this project is that racial/ethnic and family income groups in local areas have the same prevalence rates as those groups at the state or regional level.

A more complex approach to synthetic estimation involves regression techniques, whereby the data for the parent geography are used to estimate a regression equation with multiple independent variables. The resulting equations can be used to predict prevalence rates for smaller geographic areas based on the specific array of independent variables that are present in a local area. However, regression-based synthetic estimation is complicated by the small samples (and large sampling errors) typically associated with local area data. Moreover, regression-based synthetic estimation models introduce a temporal challenge. Users essentially have a handful of options:

- Use the base-year model, unaltered, for all subsequent years;
- Re-specify the model for each year; or
- Pool data across all available years to produce a multi-year model (the model can be used to produce either single-year or multi-year estimates).

The base-year model approach is problematic because covariates from the base year may not be optimal in subsequent years. Re-specifying the model optimizes the results for any given year, but introduces problems for time series analysis because "large fluctuations in the selected covariates raise questions about the stability of the small area predictions" and the ability to monitor trends over time.[5]

Re-specification of the regression model reduces model stability and limits the comparability of the estimates over time. Pooling data to create a model that is strong across multiple survey years may introduce other challenges as the model "does not discriminate between temporal and cross-sectional effects."[6]

In contrast, the ratio-based synthetic estimation approach allows a user to apply the same model framework for any given year. An update to the model requires only an update to the coefficients, not to the basic model structure (i.e. variables included in the model). This facilitates time-series analysis, which is an important consideration for this project. This type of model has been used, successfully, to estimate health measures for a wide variety of conditions including, but not limited to, disability status, contraceptive use, household sanitation, prenatal care, and post-natal outcomes.[7,8]

We chose the ratio-based method because it is easier for nontechnical data users to understand and replicate, and it ensures the comparability of the data over time.

Synthetic estimates were produced for 26 key indicators of child well-being from the 2011-2012 NSCH microdata file. Key indicators were selected with input from Child and Adolescent Health Measurement Initiative (CAHMI) and CityMatCH staff. In this paper, we report on key findings based on one of these 26 measures: childhood obesity/overweight. We chose to focus on obesity because the number of children at risk of being overweight or obese has increased dramatically in recent years, posing a major, long-term health risk to children.

ACS population data were cross tabulated by four race/ethnic groups (matching the race4_11 variable in the NSCH), four ratios of family income to the FPL (matching the povlev4_11 variable in the NSCH), and three age groups (ages 0 to 5, 6 to 11, and 12 to 17). ACS data were rounded and suppressed in accordance with Census Bureau nondisclosure rules.

A local estimate will deviate from the state estimate if two conditions are present: 1) the racial/ethnic and/or income distribution in the local area differs from that of the state; and 2) prevalence rates vary across different racial/ethnic and income groups.

In our analysis, NSCH prevalence rates for a parent geography (state or census region, described in more detail below), broken down by race/ethnicity and family income level, were applied to population data at the city or county level. The resulting estimates were then summed to the city or county total and divided by the population for that target geography (see Figure 1). In essence, the ACS population totals by race/ethnicity and income are used to reweight the NSCH rates to produce prevalence rates at the city and county levels.

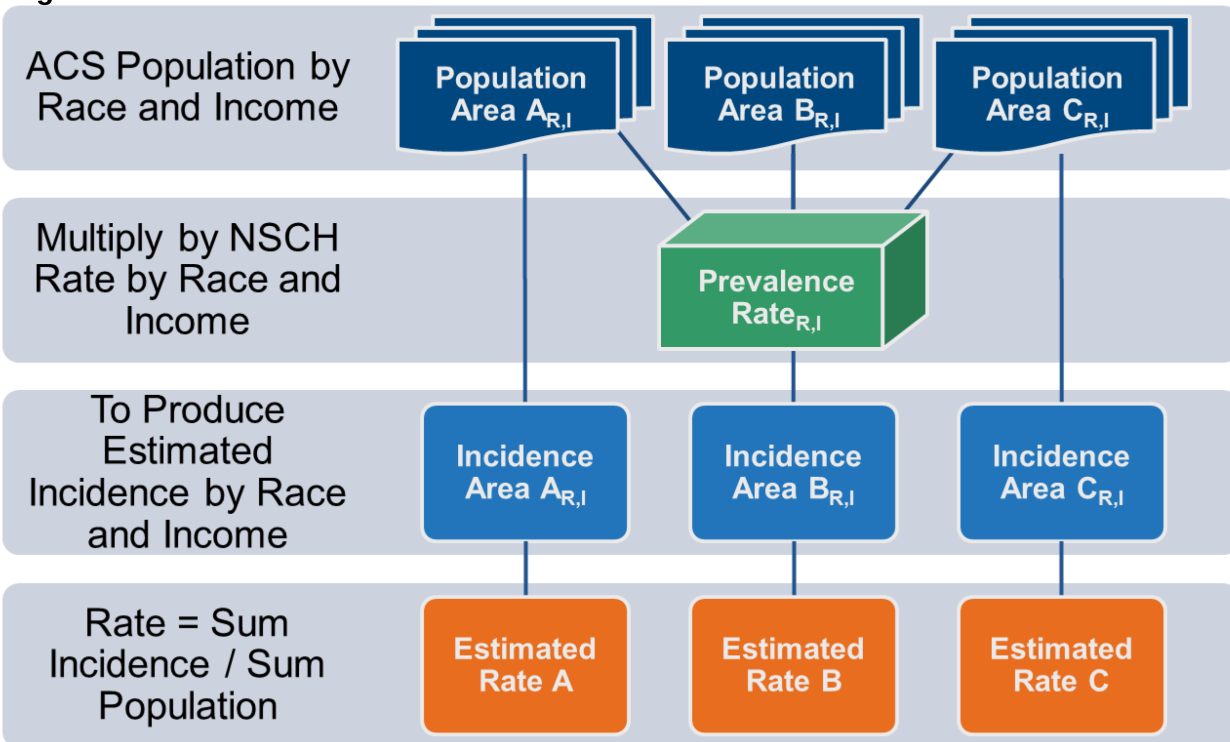
This process is represented by the following equation:

$$Synthetic\ Prevalence\ Rate\ = \sum_{for\ each\ r,i} \frac{Pop_{r,i}}{Pop}\, x\ Rate_{r,i}$$

Where:

Synthetic Prevalence Rate = synthetic prevalence rate estimate for a target geography
Pop = American Community Survey population data for the target geography
Rate = Prevalence rate calculated from the NSCH for the parent geography
r = each racial category
i = each income category

Figure1: Estimation Process



Population universe

For most of the indicators, data are for the population ages 0 to 17, ages 0 to 5, or ages 6 to 17. However, some NSCH variables are reported for other age groups. In these cases, it was not possible for us to collect the data directly from the Census Bureau because of the risk of disclosing information about very small population groups. (For example, it would be possible to estimate the number of 10-to-11-year-olds by subtracting the number of 10-to-17-year-olds from the number of 12-to-17-year-olds.) Instead, we estimated the population ages 10 to 17 based on the population data for children ages 6 to 11 and ages 12 to 17. We assumed that the county and city populations were evenly distributed across the 6-to-11 age group in making these calculations.

Parent geography

The parent geography for a given city or county is its state, unless the state-level NSCH prevalence rate is derived from fewer than 20 unweighted cases. When NSCH rates are based on fewer than 20 cases at the state level, the model draws the relevant rate from the broader census region (Northeast, Midwest, South, or West).

For example, when calculating the synthetic prevalence rate of preventive medical visits for cities and counties in Alabama, there are a sufficient number of unweighted NSCH cases to use state-level prevalence rates for Hispanics at 0 to 99 percent of the Federal poverty level and at 100 to 199 percent of the poverty level. But for Hispanics at 200 to 399 percent of the poverty level, there are fewer than 20 unweighted cases for the state of Alabama, so the rate for the

South Region was used to estimate the number of children receiving preventive medical visits for that population subgroup. The totals were then summed across all population subgroups for the target geography, regardless of whether the prevalence rate was derived from the state or regional level.

Geographic Areas

PRB compiled ACS population data and synthetic prevalence rate estimates for cities and counties if their total population in the 2010-2012 American Community Survey was at least 100,000. However, there were three college towns that met the population threshold but where there were no children under age 18:

- South Bend, Indiana (adjacent to Notre Dame University);
- Edison township, New Jersey (adjacent to Rutgers University); and
- Murfreesboro, Tennessee (home to Middle Tennessee State University).

In addition, Wyoming is not represented in the custom ACS tabulation because there were no counties or cities in the state that met the 100,000 population threshold in 2012. To ensure some coverage in this state, a second tabulation, run on the ACS Public Use Microdata File, was produced to estimate the population by age, race/ethnicity, and ratio of family income to poverty level for the combined counties of Albany and Laramie in Wyoming.

EVALUATING THE RESULTS

Table 3 presents synthetic estimates of the number of overweight/obese children ages 10 to 17 in Baltimore, Maryland. Obesity estimates were calculated by multiplying the NSCH prevalence rates from Table 1 by the ACS population data presented in Table 2.

Table 3: Estimated Number of Overweight/Obese Children in Baltimore

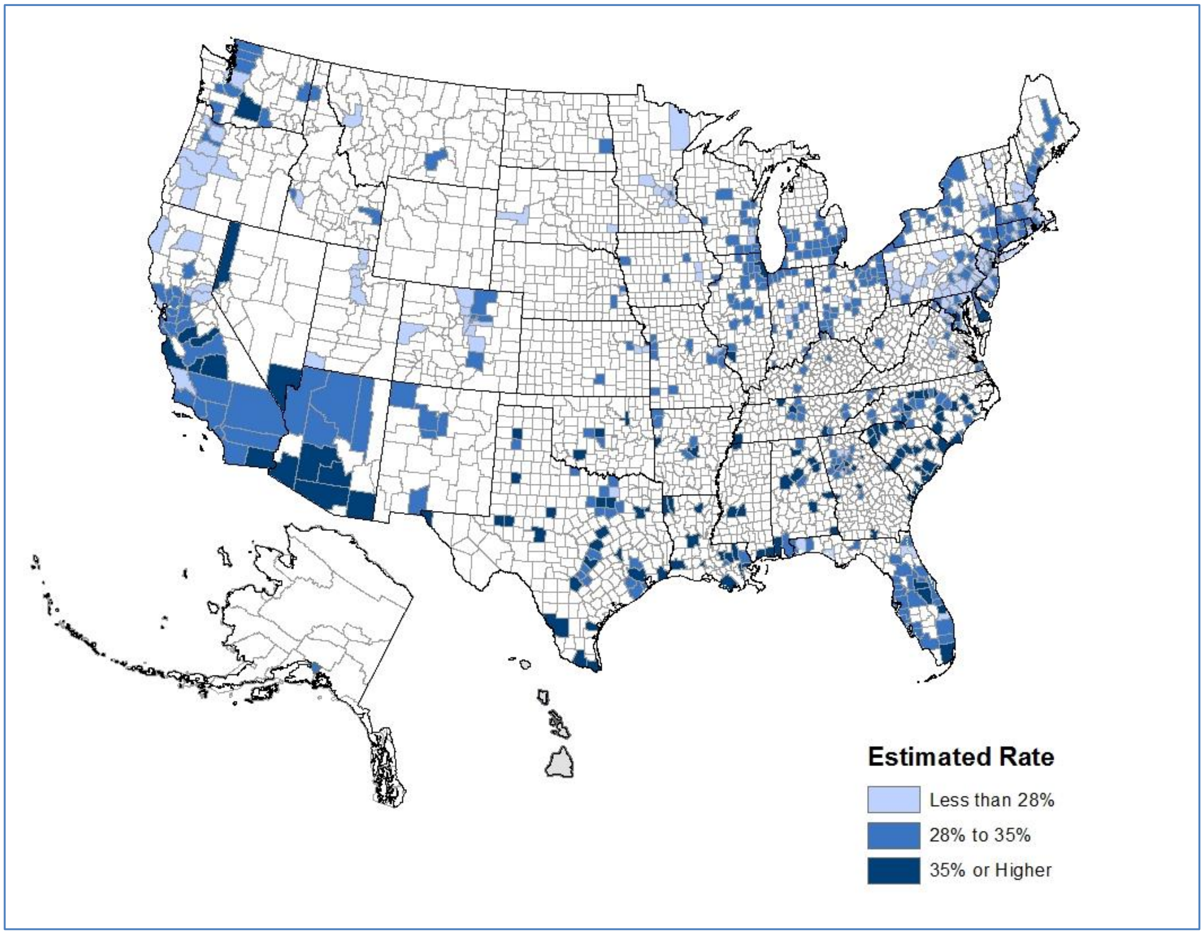
Est. Number of Overweight /Obese BALTIMORE	TOTAL	0-99% FPL	100-199% FPL	200-399% FPL	400% FPL or Higher
TOTAL	**26,564**	9,787	10,149	4,782	1,846
Hispanic	850	388	212	199	51
White, non-Hispanic	2,046	531	439	566	520
Black, non-Hispanic	22,775	8,600	9,253	3,756	1,166
Other, non-Hispanic	883	268	245	261	109

Source: PRB.

The estimated number of overweight/obese children in Baltimore City, MD (26,564) is divided by the estimated population ages 10 to 17 (54,028) to create a total overweight/obesity rate for children ages 10 to 17 in Baltimore (49 percent). About 86 percent of overweight/obese children in Baltimore are African American (22,775 out of 26,564).

Figure 2 shows patterns of childhood overweight/obesity across large counties in the United States. Overweight/obesity rates are generally lower in counties in the northern states and higher in the South. There are also clear regional patterns within some states. For example, overweight/obesity rates are relatively low in northern California compared with rates in the southern part of the state. There are also many metropolitan areas—including Atlanta, Baltimore, Chicago, Dallas, Detroit, New York, Richmond, and Washington DC—where overweight/obesity rates are relatively high in central city areas and much lower in the surrounding suburban counties. Explaining these regional patterns is outside the scope of this analysis, but variations in overweight/obesity rates have been linked to a broad combination of environmental factors—including climate, land use, population density, and cultural determinants—as well as individual and family characteristics such as gender, age, race/ethnicity, and socioeconomic status. [9,10]

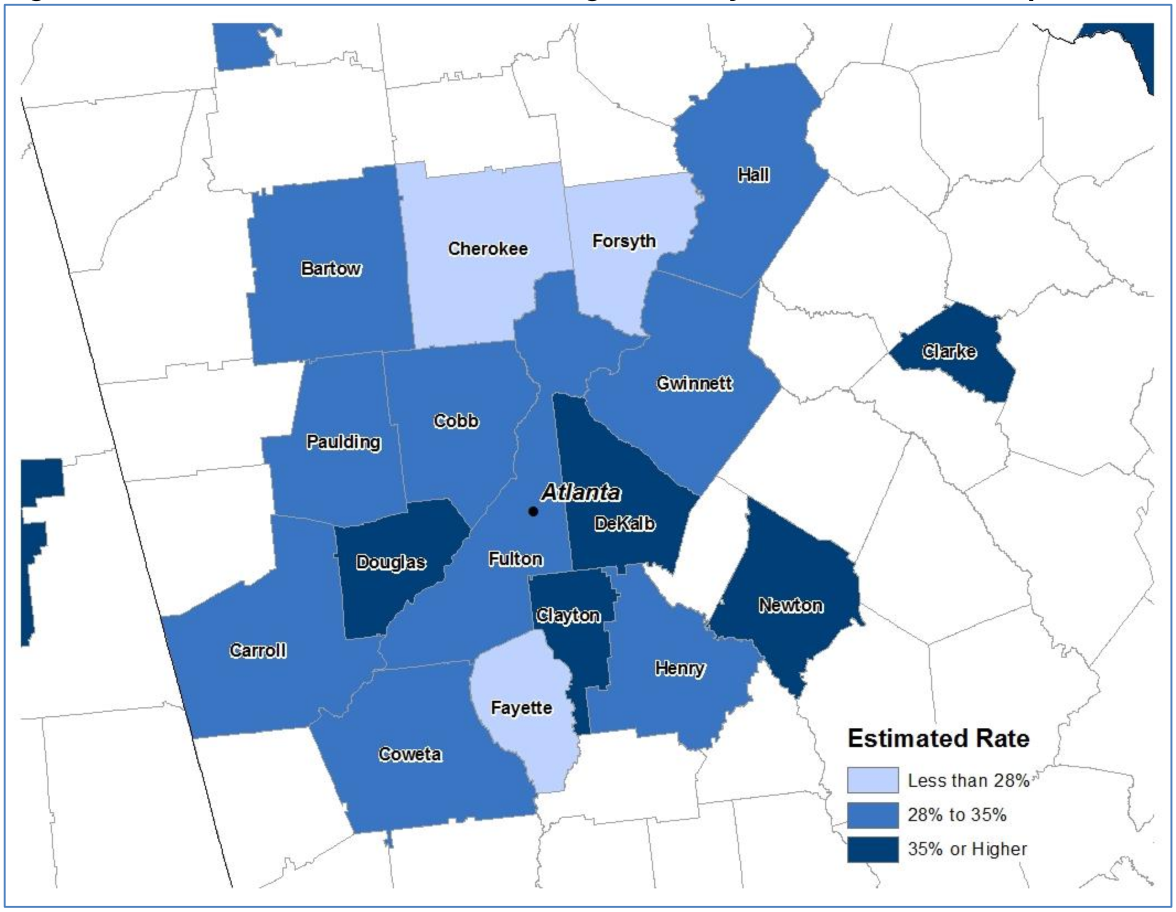
Figure 2: Estimates of Childhood Overweight/Obesity for Large Counties in the United States



Source: PRB.

The Atlanta Metropolitan Area shows a clear pattern with higher overweight/obesity rates in Clayton and DeKalb Counties—both majority African American counties near the city center—and lower rates in majority-white Forsyth County (see Figure 3).

Figure 3: Estimates of Childhood Overweight/Obesity in the Atlanta Metropolitan Area



Source: PRB.

There are several sources of potential error in these synthetic estimates:

- Sampling error in the NSCH;
- Sampling error in the ACS data affecting the population reweighting scheme;
- Rounding error in the ACS data affecting the population reweighting scheme; and
- Variation in the health and well-being of children across different racial/ethnic and family income groups at the state and local level.

We evaluated the results of our analysis using two different methods: 1) Testing synthetic estimation procedures at the state level; and 2) Measuring sampling error in the data.

Testing Synthetic Estimation at the State Level

We tested the validity of the synthetic estimation method by applying NSCH prevalence rates by race/ethnicity and family income at the regional level (Northeast, Midwest, South, West) to state-level ACS estimates of the child population by race/ethnicity and income. We then compared the resulting synthetic state prevalence rates with published NSCH rates to see how well the model performed.

The accuracy of our synthetic estimates is evaluated based on two measures: 1) Mean Absolute Percent Error (MAPE) which indicates precision of the synthetic estimates when compared with actual state values; and 2) Mean Algebraic Percent Error (MALPE) which indicates positive or negative bias of the synthetic estimates when compared with actual state values (see Table 4). Across the 50 states and the District of Columbia, the MALPE score was 1.5 percent, while the MAPE score was 7.5 percent. In one-fourth of the states, the difference between the synthetic and published estimates was less than 1 percentage point. For the majority of states, the estimated state values were within 2.5 percentage points of the published NSCH values.

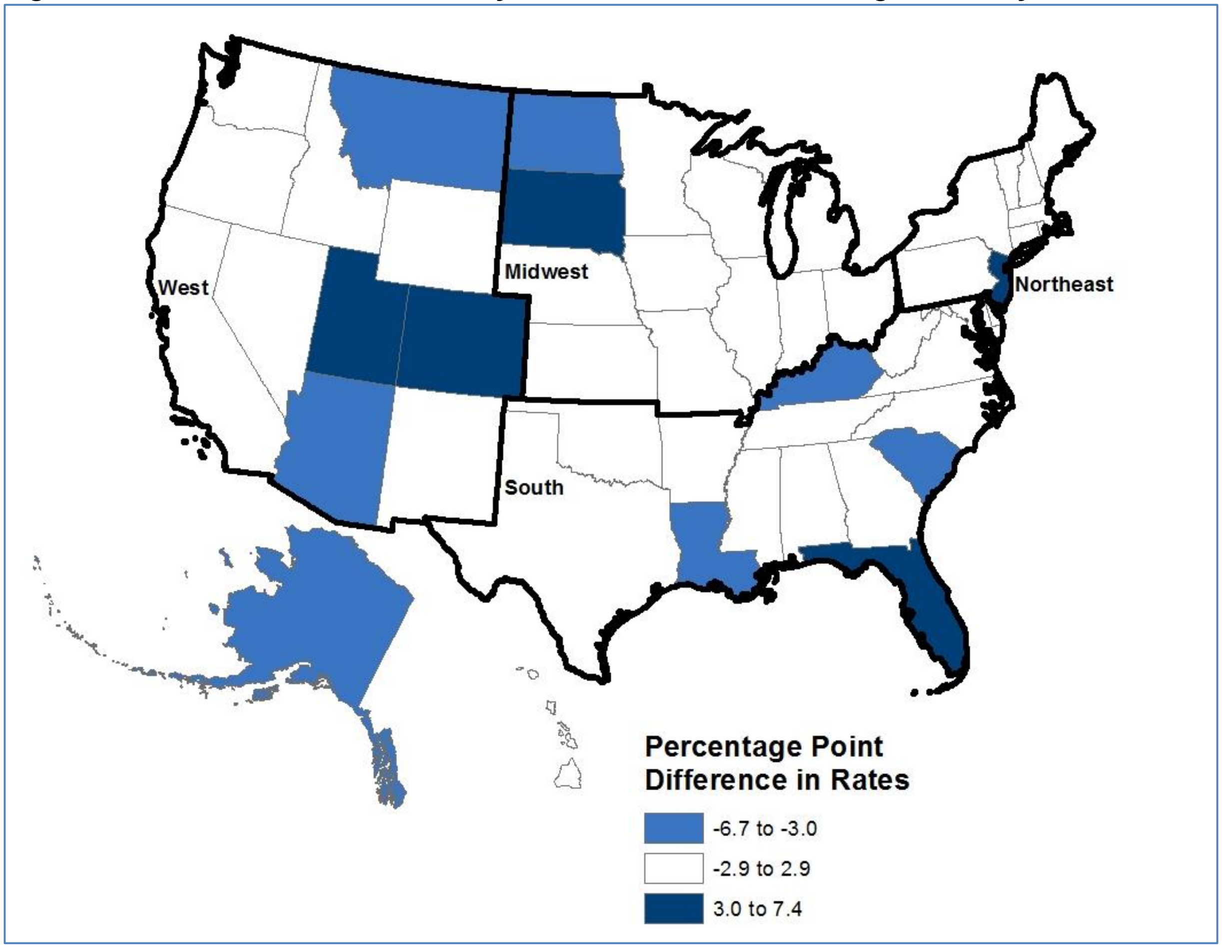
Table 4: Summary of State Comparisons

	Overweight & Obesity
U.S. Rate	31.3
Range of state values	17.3
High (MS)	39.7
Low (UT)	22.4
MAPE	7.5
MALPE	1.5
Max. underestimate	-6.7
Max. overestimate	7.4
Number of states within 1 pt.	13
Percent within 1 pt.	25%

Source: PRB

There were no clear regional patterns in the difference between the synthetic estimates and published data. There were two states—Arizona and North Dakota—where NSCH published estimates were more than 5 percentage points higher than the synthetic estimates (see Figure 4). In three other states—Colorado, Florida, and New Jersey—published estimates were more than 5 percentage-points lower than our synthetic estimates.

Figure 4: Difference Between Survey and Estimate for Overweight/Obesity



Source: PRB.

Some of these errors reflect the unique health profiles of certain states relative to the broader census region in which they are located. For example, the percent of children who are overweight or obese in Colorado (23 percent) is much lower than that of children in California (31 percent). This difference is important because the health status of different racial/ethnic and income groups in California factor highly in the prevalence rates for the West Region, in which Colorado is located. North Dakota provides a similar example, where the published rate of children who are overweight or obese (36 percent) is much higher than that of other states in the Midwest.

We also tested the accuracy of our synthetic state prevalence rates using different combinations of racial/ethnic and income categories, i.e., more detailed or less detailed categories for income and for race/ethnicity. The results of this analysis showed that our model, based on 16 detailed racial/ethnic and income categories, yielded better estimates (lower MAPE and MALPE scores) than a model that used fewer categories. Including more than 16 racial/ethnic categories or adding additional indicators to the model could potentially yield more accurate results, but the NSCH and three-year ACS samples would not support more detailed crosstabulations.

Measuring Sampling Error

We used Coefficients of Variation (CVs) to determine the reliability of ACS and NSCH estimates. CVs are calculated using the following formula:

$$CV = \frac{SE_{gri}}{x_{gri}} * 100$$

In this formula, the CV for each geographic area (g), racial/ethnic group (r), and income group (i) was calculated by dividing the standard error (SE) of the estimate by the estimate itself and multiplying the result by 100. Estimates with small CVs are considered more reliable than estimates with larger CVs. The Centers for Disease Control and Prevention reviewed the criteria for data suppression used by 22 of the 23 major data systems and found that 30 percent is a common CV level used to suppress estimates for many data systems. [11] However, the Census Bureau suppresses tables based on the three-year ACS if the median CV for cells in the table is greater than 61 percent.

We evaluated our synthetic state-level prevalence rates against published NSCH prevalence rates using 30 percent and 60 percent cutoffs for CVs. We found that at the state level, very few population groups would be suppressed at either at the 30 percent or 60 percent level. Therefore, the MAPE and MALPE scores based on the two data suppression methods were very similar.

Following the Census Bureau's general guidelines, we suppressed ACS counts of the child population by race/ethnicity and income if the coefficient of variation (CV) for the cell estimate was greater than 60 percent. We also suppressed cells in the race/ethnicity and income matrix if the corresponding NSCH prevalence rate for that group had a CV that was greater than 60 percent. However, most CVs for the NSCH prevalence rates were relatively small, since we had already incorporated a method to improve the reliability of the NSCH prevalence rates (substituting regional rates for state rates in cases where there were fewer than 20 unweighted respondents for a given racial/ethnic and income group.)

Ideally, we would provide standard errors associated with our synthetic county- and city-level prevalence rates. However, direct calculation of standard errors for individual synthetic estimates is not feasible.[12] Calculating the net error associated with the ACS reweighting scheme is also complicated because the Census Bureau's published formulas do not function well in cases where multiple ACS estimates need to be combined to produce a derived estimate.[13] In these cases, the calculated standard error may overestimate or underestimate the actual standard error, depending on whether the estimates that are being combined are positively or negatively correlated.

Comparison with Direct Estimates for Local Areas

We had hoped to compare our synthetic estimates for a few large counties with direct estimates from the NSCH. However, there is a potential issue in applying NSCH prevalence rates that are designed to be representative at the national and state levels to smaller geographic areas, so the accuracy of these direct local estimates is unclear. For more information about this issue, see Lee Mobley's editorial in Spatial Demography.[14]

DISCUSSION

Lack of reliable data at the local level leaves public health agencies with a gap in knowledge. This project demonstrates how state prevalence rates can be combined with local population estimates to produce prevalence rates for local areas. We hope that these estimates will aid local policymakers in developing targeted programs and policies to improve children's health in their community.

One of the main strengths of synthetic estimation is that it is easily explained and easy to understand, compared with more complex regression-based methods. Population estimates are also readily available from the U.S. Census Bureau and can be updated on a regular basis. Another strength of synthetic estimation is that local estimates tend to be closely aligned with the state-level estimates on which they are based.

However, there are two major limitations of this analysis. First, sampling error in the state-level prevalence rates and local population estimates may be compounded in the final synthetic estimates, resulting in erroneous conclusions about the well-being of children in local areas. Second, and more importantly, the model assumes that prevalence rates for people in different racial/ethnic and income groups in local areas match those at the state level. In other words, it is assumed that in a given state, race/ethnicity and family income are the sole factors contributing to geographic differences in the well-being of children. This may or may not be the case, and there may be interaction effects between race/ethnicity and income at play.

As a next step, we plan to incorporate regression models to estimate the relationships between race and income and several key dependent variables in the NSCH microdata file. These models would indicate which variable—race or income—and which categories of these variables have stronger independent associations with children's well-being, and whether there are any interactions between the two variables. For example, does the effect of income on children's health vary for different racial ethnic groups, or does the association between race and child well-being differ for families with varying income levels? Understanding these micro-level relationships will help us develop more robust models to predict county-level estimates of child well-being.

We would also like to explore the possibility of introducing additional variables into the model to improve the accuracy of the synthetic estimates. Prior research has also shown that there are many other factors besides race/ethnicity and income that could affect small-area health prevalence rates, including urban/rural residence, access to health care, family structure, environmental factors, behavioral factors, social capital, and neighborhood safety and support.

Endnotes

[1] From 2005 through 2010, about 2.9 million addresses were sampled annually, resulting in about 1.9 million final interviews. However, in 2011 and going forward, the annual sample size was increased to 3.5 million addresses. Because the sample size increase in 2011 did not take effect until June, the sample size for the 2011 ACS was about 3.3 million addresses, resulting in approximately 2.1 million interviews.

[2] In this paper, the family income of children is measured as a percentage of the Federal Poverty Level (FPL). However, for ease of reference, we call it simply "family income" in the text from this point forward.

[3] National Center for Health Statistics, *Synthetic State Estimates of Disability*, Public Health Service, PHS Publication No. 1759, Washington: U.S. Government Printing Office, 1968.

[4] Child and Adolescent Health Measurement Initiative, "Local Uses of National and State Data," http://childhealthdata.org/docs/nsch-docs/local-use-of-state-data-and-synthetic-estimates.pdf.

[5] Bart Buelens and Jan van den Brakel. (2014), "Model selection for small area estimation in repeated surveys," accessed online at http://www.cbs.nl/NR/rdonlyres/308ED398-714A-41A4-A57C-9DCCC3F30D35/0/201423x10pub.pdf

[6] Bart Buelens and Jan van den Brakel. (2014), "Model selection for small area estimation in repeated surveys," accessed online at http://www.cbs.nl/NR/rdonlyres/308ED398-714A-41A4-A57C-9DCCC3F30D35/0/201423x10pub.pdf

[7] Tsukasa Namekata, Paul S. Levy, and Thomas W. O'Rourke. (1975), "Synthetic estimates of work loss disability for each state and the District of Columbia," accessed online at http://www.ncbi.nlm.nih.gov/pmc/articles/PMC1438269/.

[8] Ahmet S. Turkyilmaz. (2003), "Application of Synthetic Estimation to Selected Demographic and Health Indicators as a Traditional Small Area Estimation for Provinces of Turkey," accessed online at http://www.hips.hacettepe.edu.tr/nbd_cilt25/turkyilmaz.pdf.

[9] Michimi A. and MC Wimberly, (2010), "Spatial patterns of obesity and associated risk factors in the conterminous U.S." *American Journal of Preventive Medicine* 39 (2):e1-12.

[10] Wang, Y. and M.A. Beydoun (2007), "The Obesity Epidemic in the United States—Gender, Age, Socioeconomic, Racial/Ethnic, and Geographic Characteristics: A Systematic Review and Meta-Regression Analysis," *Epidemiological Review* 29 (1): 6-28.

[11] Richard J. Klein et al. (2002), "Healthy People 2010 Criteria for Data Suppression," accessed online at http://www.cdc.gov/nchs/data/statnt/statnt24.pdf.

[12] For more information, see M.E. Gonzalez (1973), "Use and evaluation of synthetic estimators," In *Proceedings of the Social Statistics Section*. 33-36. American Statistical Association, Washington, DC.

[13] U.S. Census Bureau, American Community Survey Office (2013), "American Community Survey Multiyear Accuracy of the Data," accessed online at www.census.gov/acs/www/Downloads/data_documentation/Accuracy/MultiyearACSAccuracyofData2012.pdf.

[14] Lee R. Mobley (2014), "Drawing Over-reaching Conclusions from Spatial Health Data," *Spatial Demography* 2(1): 66-71, accessed online at http://spatialdemography.org.