

**Supporting Statement B**

**for**

**The Healthy Communities Study:  
How Communities Shape Children's Health (NHLBI)**

**July 11, 2013**

Name: Sonia Arteaga  
Address: 6701 Rockledge Drive, Suite 10018, Bethesda, Maryland 20892  
Telephone: (301) 435-6677  
FAX: (301) 480-5158  
Email: [arteagass@nhlbi.nih.gov](mailto:arteagass@nhlbi.nih.gov)

**TABLE OF CONTENTS**

B. COLLECTIONS OF INFORMATION EMPLOYING STATISTICAL METHODS.....1

    B.1 Respondent Universe and Sampling Methods.....1

        B.1.a Design Summary.....1

        B.1.b Respondent Universe – Communities.....1

        B.1.c Respondent Universe – Schools and Child/Parent Participants.....12

        B.1.d Respondent Universe – Key Informants.....15

        B.1.e Power Calculations.....17

    B.2 Procedures for the Collection of Information.....22

        B.2.a Community-Level Data Collection.....22

        B.2.b Data Collected at the Child/Parent Participant Level.....26

    B.3 Methods for Maximizing Response Rates.....29

    B.4 Tests of Procedures or Methods to be Undertaken.....30

    B.5 Training and QA/QC Methods.....31

    B.6 Consultations and the Project Team.....31

**LIST OF TABLES**

Table B.1.1 The Largest 15 Counties in the United States (as per 2010 U.S. Census).....7

Table B.1.2 Number of Planned Child/Parent Assessments by Type of Community.....14

Table B.1.3 Recruitment Goals for Children by Grade and Gender for Non-Minority  
Communities.....15

Table B.1.4 Recruitment Goals for Children by Grade and Gender for Minority Communities  
.....15

Table B.1.5 Priority/Alternate Community Key Informants Targeted for the HCS.....16

Table B.1.6 Estimated % Difference in BMI that Can Be Detected for a One-unit Difference in  
Intensity Score Assuming a Cross-sectional Analysis with 80 % Power.....18

Table B.1.7 Estimated Power to Detect a 3% Change in BMI for a One-unit Difference in  
Intensity Score Assuming a Longitudinal Analysis.....19

Table B.1.8 Estimated Percent Difference for Binary Diet or Physical Activity Measure for a  
One-unit Difference in Intensity Score Assuming a Cross-sectional Analysis with  
80 % Power.....20

**LIST OF FIGURES**

Figure B.2.1 HCS Primary Data Collection Activities in Wave 2, 2013-2015.....22

## **LIST OF ATTACHMENTS**

- Attachment 1. Membership of the Certainty Community Selection Committee**
- Attachment 2. Nomination Form for Community Selection**
- Attachment 3. Assessment of Statistical Power for the HCS**

## **B. COLLECTIONS OF INFORMATION EMPLOYING STATISTICAL METHODS**

### **B.1 Respondent Universe and Sampling Methods**

#### ***B.1.a Design Summary***

The Healthy Communities Study (HCS) will collect both current/cross-sectional and retrospective data on participants and their parents in 264 communities across the United States during Wave 2. The HCS will use a three-stage sample design to recruit elementary and middle-school children. The first stage will use a hybrid design to sample communities that includes a stratified national probability-based sample (NPBS) of communities combined with a sample of communities selected with certainty that have promising programs and policies aimed at reducing childhood obesity. The process for selecting the sample of communities was completed in August 2012, with the census tracts for the National Probability Based Sample drawn, and the certainty communities selected. In the second-stage, schools from within selected communities will be chosen based on location, demographic characteristics, and interest in participation. In the third-stage, children/parents will be recruited based on grade, gender race/ethnicity, and their family's interest in participating. A limitation of recruiting participants based upon interest is that the findings will not be representative of students across the country and thus findings will not be generalizable to all U.S. public elementary and middle school children.

#### ***B.1.b Selection of Communities***

In the HCS, a community is defined as the catchment area of a local public high school. The HCS will use a hybrid design for selecting communities in Wave 2. The NPBS used a stratified sampling approach with probability proportional to the number of children aged 3-15 years to select 195 census tracts across the continental US. The strata were based on various factors including geographic region, income, race/ethnicity, urbanicity, and population size of the county. The purpose of the 195 NPBS is to ensure that the HCS provides representation of communities across the US.

A certainty community selection committee (CCSC) (please see **Attachment 1** for a list of CCSC members) independently identified 86 areas to ensure the inclusion of communities with promising programs and policies aimed at reducing childhood obesity. Within each of the 86 geographic areas selected by the committee, a census tract was selected with probability proportional to size (similar to the NPBS).

The 281 sampled census tracts were then used to identify communities for the HCS through identification of the closest public high school to the centroid of each selected tract. In some cases, there were multiple sampled census tracts that pointed to the same public high school which reduced the total number of Wave 2 communities from the originally proposed 275 communities to 264 communities.

The following subsections provide detail on the selection of Wave 2 communities into the HCS.

##### ***B.1.b.1 Processes for Gathering Information for Community Selection and Stratification***

The project team developed and populated a comprehensive and hierarchical Geographical Information Systems (GIS) database of states, counties, municipalities, and neighborhoods that

have prevention policies and programs targeting childhood obesity. A hierarchical structure was needed because programs and policies are implemented at differing levels of geographic specificity, most of which will not be consistent with the high-school catchment area definition of a community that is being used in this study. GIS software was used to associate census tracts, school districts, and high-school catchment areas with broader geographic boundaries and vice versa. The database was used for the selection of certainty communities, the stratification of counties in the design, and to leverage an existing GIS database (Advancing the Movement) developed with support from the National Collaborative on Childhood Obesity Research (NCCOR).

The GIS database was used to develop a pre-selection scoring of communities and counties. Variables that contributed to this pre-selection scoring include the following:

- Whether the program or policy was developed and implemented at the community, county or state/regional level
- Number of programs and policies identified from listings provided by NHLBI, RWJ, YMCA, and others
- Selected variables from CDC Children’s Food Environment State Indicator Report
- Selected variables from CDC State Indicator Report on Physical Activity
- State sales taxes levied on soda sold in retail stores and vending machines and on chips/pretzels sold in vending machines
- Selected nutrition variables at the local level obtained from the CARES database
- Classification of Laws Associated with School Students (C.L.A.S.S.) scoring of state-level codified laws for physical education (PE) and nutrition in schools
- Number of independent nominations received by members of the HCS Steering Committee, state health officials, and representatives from non-federal funders

To populate the GIS database multiple approaches were employed while using a standardized information-gathering template as shown in **Attachment 2**. This template was used to gather pre-selection information on programs and policies that occur at different levels of geographic specificity (state, county, municipality, or neighborhood). The sources of data that were used to populate the GIS database are listed below:

1. **HCS Steering Committee Members:** We included recommendations/nominations of specific communities to be considered as candidates for certainty communities from members of the HCS Steering Committee (or their designees). The person who nominated a community provided information using the template shown in **Attachment 2**. A variable was integrated into the GIS database that captures the fact that a particular community was recommended by one or more members of the Steering Committee – one of the factors that is used to stratify communities as described below in Section B.1.b.2.
2. **Federal Funders:** We performed a census of the large federally supported programs, such as the National Heart, Lung, and Blood Institute’s (NHLBI’s) **We Can! (Ways to Enhance Children’s Activity & Nutrition)**® program, and the Centers for Disease Control and Prevention’s (CDC’s) Communities Putting Prevention to Work (CPPW) program, to gather information on program characteristics (i.e., community name, level of geographic specificity

funded, level of funding, duration of funding, and general information about the program goals).

3. ***Non-federal Funders:*** We contacted members of the Convergence Partnership, a collaboration of large funders of childhood obesity programs, to gather data on programs they fund. The Convergence Partnership includes the California Endowment, Kaiser Permanente, the Kresge Foundation, Nemours, the Kellogg Foundation, and CDC. A maximum of nine non-federal entities were contacted due to Office of Management and Budget (OMB) restrictions in data gathering. We asked program funders to complete the information-gathering template for communities that were funded to reduce childhood obesity over the past decade (i.e., community name, level of geographic specificity funded, level of funding, duration of funding and general information about the program goals).
4. ***Online Policy Databases:*** We researched national, state and local policies that affect childhood obesity, such as those adopted by the state of Arkansas to promote BMI evaluation and appropriate nutritional and physical activity follow-up among school-aged children. There are multiple documented public databases available, which provide information about such policies, including nutritional and physical activity policies in schools (e.g., the National Cancer Institute's (NCI's) Classification of Laws Associated with School Students database). These policies were also incorporated into the GIS database, including duration of policy and level of geographic specificity (school district, town/city, county, state) where available.
5. ***State Health Departments:*** Following OMB approval of the original protocol (OMB Notice of Approval 0649 dated 1/30/2012), we surveyed 28 leaders involved in combating childhood obesity within various states (e.g., the head of the Chronic Disease Prevention program within the State Health Department). For each community state leaders identified as having promising programs and policies that address obesity for children and youth, we asked for the program names, funding duration and amount, and funding organization.
6. ***General Program/Policy Information:*** In data gathering (before and after OMB approval of the original protocol), we conducted initial outreach via e-mail to set up telephone appointments with funders/sponsors of programs and policies (e.g., YMCA) to gather general information, key documents (via e-mail, mail, links to Websites, etc.), funding levels, and duration of funding within specific geographic areas via databases. We asked these organizations to provide us with available information regarding their programs and policies, but did not ask them to complete any sort of questionnaire. Thus, there was no participant burden from these individuals.

#### ***B.1.b.2 Selection of Probability-Based Sample of Communities***

We selected the first 195 areas to be considered for Wave 2 using a stratified probability-based approach. Following this approach, we grouped census tracts into strata based on race/ethnicity, urbanicity, income, region of the county, and pre-selection score (methods used to establish this pre-selection scoring are described above in Section B.1.b.1). Each census tract across the United States was assigned to one stratum. Within each stratum, a single census tract was selected using a probability proportional to size algorithm (where size refers to the number of children aged 3-15 years in each census tract based on Census 2010 data). After each census

tract was selected, our team determined which high school was in closest proximity to the selected census tract to identify the specific community for inclusion into the study.

It should be noted that the research team investigated multiple options for drawing the sample, with a strong preference of stratifying the nation based on high-school attendance boundaries. Unfortunately these boundaries change frequently based on district needs (e.g., efficient assignment of students to different schools within a district) and there is no current universal data source that captures this information systematically across the entire US. Therefore, it would be difficult to develop an appropriate sampling frame for the study based on high-school catchment area. We considered a two-stage sampling scheme with sampling at the first stage at the county or school district level, followed by identification of high-school catchment areas within selected counties or district at the second stage of sampling. Our investigation into this method revealed that the level of specificity in our stratification variables (race/ethnicity, urbanicity and income) became diluted within these larger areas (counties and districts), leading to suboptimal stratification.

Therefore, the stratified sampling of census tracts was selected as the best option for drawing an appropriate sample of communities into the study. This method of selecting census tracts and communities within counties helps ensure all communities across the United States have a known probability of being included in the HCS. Construction of the strata ensures representation across a range of geographically, demographically, and culturally diverse communities.

The stratification variables used to classify census tracts are detailed further below:

1. Region of the Country

Region	States included
1 (Northeast)	CT, MA, ME, NH, NJ, NY, PA, RI, VT
2 (Midwest)	IA, IL, IN, KS, MI, MO, MN, ND, NE, OH, SD, WI
3 (South)	AR, DE, DC, AL, FL, GA, KY, LA, MD, MS, NC, OK, TN, TX, SC, VA, WV
4 (West)	AK, CO, AZ, CA, HI, ID, MT, NM, NV, OR, UT, WA, WY

2. Urban / Suburban / Rural Designation

The methodology we are utilizing for urban/suburban/rural designation is based on the Rural-Urban Commuting Area (RUCA) system which was developed by the U.S. Departments of Agriculture and Health and Human Services and a multi-state/university collaborative in the 1990s. This system uses commuting patterns to distinguish between urban, suburban, large rural, and small rural areas (vs. the more simplistic urban/rural designation provided by the Census). The RUCA system classifies each census tract based on the following characteristics:

- **Urban Core** areas are contiguous, built-up areas containing 50,000+ people – these correspond to the Census Bureau’s Urbanized Areas.
- **Suburban** areas have high commuting flows to Urban Core areas. These include areas in which 30-49% of the population commutes to Urban Core areas for work.

- **Large Rural Towns** have a population between 10,000 and 49,999 people and include surrounding rural areas with at least 10% primary commuting flows to these towns, as well as secondary commuting flows of 10% or more to Urban Core areas.
- **Small Towns and Isolated Rural Areas** include towns with populations <10,000 people and their surrounding areas, as well as isolated rural zones more than 1 hour (by car) away from the nearest city.

For the purposes of the HCS, we have collapsed Large Rural Towns with Small Towns/Isolated Rural Areas to form a single “Rural” category. The RUCA system applies to 2000 census tract boundaries (they have not been updated for 2010). Given this limitation, we are mapping all of the Census 2010 data (for demographic characteristics) into the Census 2000 boundaries (there are standard GIS tools for doing this translation between Census 2000 and 2010 boundaries).

### 3. Race/Ethnicity

The HCS is designed to specifically over sample communities with high proportions of African American residents and Hispanic/Latino residents, as these populations are at increased risk for childhood obesity.

Stratification of census tracts for oversampling these important minority populations is based on the following definitions:

- A census tract will be stratified into a high-proportion of **African American** group if 30% or more of the total population in the census tract is African American\*
- A census tract will be stratified into a high-proportion of **Hispanic/Latino** group if 30% or more of the total population in the census tract is Hispanic/Latino\*
- A census tract will be stratified into the **Other** group if less than 30% of the population in the census tract is African American and less than 30% of the population in the census tract is Hispanic/Latino

The selection of 30 percent was based on the following definition of a minority census tract:

*Banking regulations define minority communities when it comes to prohibiting redlining and other discriminatory lending practices. According to 12 USCS § 4502 (29), [Title 12. Banks and Banking; Chapter 46. Government Sponsored Enterprises] the term minority census tract means “a census tract that has a minority population of at least 30 percent and a median family income of less than 100 percent of the area family median income.*

\* For census tracts with both 30% African American and 30% Hispanic/Latino; the tract is classified in the category with the higher proportion (e.g., if a census tract contains 43% African American and 34% Hispanic/Latino residents – it would be classified as a high-proportion African American census tract).

Additionally, it should be noted that some residents within a census tract may be both African American and Hispanic/Latino. These residents would contribute to the calculation of both percentages when stratifying a census tract. Thus, the sum of the percentages (African American, Hispanic/Latino and Other) may be greater than 100% for some census tracts when doing these calculations.

#### 4. Income

The US Department of Housing and Urban Development (HUD) designates Qualified Census Tracts (QCTs) for purposes of the Low-Income Housing Tax Credit (LIHTC) program. The LIHTC program is defined in Section 42 of the Internal Revenue Code of 1986. The LIHTC is a tax incentive intended to increase the availability of affordable rental housing.

The LIHTC statute provides two criteria for QCT eligibility. A census tract must have either: 1) a poverty rate of at least 25 percent; or 2) 50 percent or more of its householders must have incomes below 60 percent of the area median household income. The area corresponds to a metropolitan or a non-metropolitan area. Further, the LIHTC statute requires that no more than 20 percent of the metropolitan area population reside within designated QCTs (This limit also applies collectively to the nonmetropolitan counties in each state). Thus, it is possible for a tract to meet one or both of the above criteria, but not be designated as a QCT.

The HCS is adopting this definition (whether a census tract qualifies for the LIHTC) as an indicator of income/poverty for the study.

#### 5. Pre-Selection Intensity rating of Community-Based Programs/Policies (High, Moderate, and Low/ None – based on a scoring procedure as described below)

Once the GIS database described in Section B.1.b.1 was populated, we developed census tract level pre-selection scores to serve as the basis for stratification. The score hierarchically integrated information across the different program and policy entries (which will occur at different levels of geographic specificity), with higher scores assigned to census tracts that had multiple program and policy entries, higher numbers of recommendations from program/policy funders or sponsors or HCS steering committee members for their promising approaches, or were located in States that had strong policy initiatives with respect top childhood obesity.

It should be noted that these scores were utilized only for the *selection* of census tracts (communities) into the study. For analyses of the HCS's results, much more rigorous data collection to assess community-based programs and policies will occur within each community selected into the study.

Scores were calculated for each U.S. census tract, and each census tract was assigned to a high-, moderate-, or low-score group based on this score - with the top 10% highest scoring counties assigned to the high-score group, the next 30% to the moderate score

group, and the lowest 60% assigned to the low score group. Our intent was to oversample communities with active programs and policies aimed at reducing childhood obesity, with approximately 65 communities being selected within each group (high-, moderate-, and low-scores).

Several counties throughout the U.S. entered our NPBS with certainty due to population size. In some counties, we selected multiple communities to preserve homogeneous sampling weights across communities selected into the HCS. The 15 largest counties in the U.S. (based on 2010 U.S. Census data) were included in the HCS, along with the number of census tracts that were selected within each of these large population centers. Table B.1.1. lists the 15 largest U.S. counties and the number of communities that were selected within each county.

**Table B.1.1 The Largest 15 Counties in the United States (as per 2010 U.S. Census)**

County	State	2010 Total Population	Number of Communities
Los Angeles County	California	9,818,605	5
Cook County	Illinois	5,194,675	3
Harris County	Texas	4,092,459	3
Maricopa County	Arizona	3,817,117	2
San Diego County	California	3,095,313	2
Orange County	California	3,010,232	2
Kings County	New York	2,504,700	2
Miami-Dade County	Florida	2,496,435	2
Dallas County	Texas	2,368,139	2
Queens County	New York	2,230,722	2
Riverside County	California	2,189,641	1
San Bernardino County	California	2,035,210	1
Clark County	Nevada	1,951,269	1
King County	Washington	1,931,249	1
Wayne County	Michigan	1,820,584	1
Total		48,556,350	30

These 15 largest counties capture approximately 16% of the U.S. population and these counties account for 30 communities; therefore, 30 of the 195 communities within the NPBS were assigned proportionally within these 15 counties as certainty counties within the NPBS.

We followed the above described stratified sampling process of selecting the remaining 165 counties. It should be noted that the number of levels of each of the stratification variables [region (4), urbanicity (3), race/ethnicity (3), income (2), and pre-selection score (3)] yield 216 unique combinations. However, there were several of these unique combinations that had either zero or very few census tracts (and children). The sampling statisticians for the HCS developed an approach for combining these unique combinations in a manner that will allow the HCS to assess differences in the association between community-based programs and policies and childhood obesity outcomes among communities with different race/ethnic profiles and among communities with different urbanicities.

Following selection of the 165 census tracts from each of the strata described above, as well as the 30 census tracts from the 15 large-population counties, we identified the high school associated with each selected census tract based on proximity.

### ***B.1.b.3 Selection of Certainty Communities***

Approximately 80 certainty communities were to be selected for inclusion in the hybrid design to ensure the sample contains communities that have a history of sufficiently promising program and policy initiatives to reduce childhood obesity.

Following selection of the 195 NPBS communities, a multi-step process was employed by the certainty community selection committee (CCSC) to select a sample of certainty communities that have the most promising programs and policies targeting childhood obesity. Please refer to **Attachment 1** for a list of the CCSC members.

Potential CCSC members were nominated by the Executive Committee using the below criteria:

- *Broad knowledge of programs and policies in childhood obesity prevention (including nutrition, PA, etc.)*
- *Experience working on these issues in the academic research community.*
- *Experience working as partners with funding agencies (e.g., the Convergence Partnership) – grantees are acceptable.*

CCSC members were selected to have a broad spectrum of experiences with community programs/policies and to be able to review the certainty community nominees and select those communities that are diverse both in terms of their programs/policies that target childhood obesity, and their demographics.

A total of 681 communities were nominated by the HCS Steering Committee, identified through members of Convergence Partnership<sup>1</sup>, interviews with non-federal sponsors of program and policies targeting childhood obesity, and interviews with state health department representatives<sup>2</sup>. These nominations were carefully reviewed by a subgroup of the Design and Analysis Subcommittee to reduce the number to a more manageable size of 267 communities; the 267 communities were selected from the following 7 sources:

1. 23 Communities Putting Prevention to Work (CPPW) sites
2. 95 sites nominated through the nomination form process (including 30 WeCan! Sites)
3. 7 Healthy Food Financing Initiative sites
4. Approximately 10 Community Transformation Grant (CTG) sites
5. Approximately 10 NY SNAP sites
6. Gold Healthy School Program sites
7. Sites nominated through the state health department interviews

---

<sup>1</sup> Members of the Convergence Partnership include Ascension Health, The California Endowment, Kaiser Permanente, Kresge Foundation, Nemours, RWJF, W. K. Kellogg Foundation

<sup>2</sup> We interviewed individuals from 29 states.

Nominated communities that did not include specific information about that particular program or policy were not included in the final database for consideration by the CCSC as they could not be evaluated to determine how “promising” the program or policy is.

The CCSC was provided with a spreadsheet with a listing of the 267 communities that included information on the community (location, community type), information on the programs/policies in that community (number of programs/policies, name, description), and the number of nominations the community had received. Each CCSC member was assigned as primary reviewer for approximately 45 communities, and a secondary reviewer for roughly another 45 communities.

The CCSC was asked to score communities with a 3, 1, or 0, using the following definitions for the scores:

3 = This community **does** warrant inclusion in the sample as one that is or recently has been highly active and promising with programs and policies to encourage healthy eating, physical activity, and weight.

1 = This community **may or may not** warrant inclusion in the sample as one that is or recently has been highly active and promising with programs and policies to encourage healthy eating, physical activity, and weight.

0 = This community **does not** warrant inclusion in the sample as one that is or has been highly active and promising with programs and policies to encourage healthy eating, physical activity, and weight.

After all the communities had been scored by the CCSC members, they were placed into a stratum, and reviewed to ensure a balance of programs and policies across the four stratification variables (ethnicity, urbanicity, income, and region) with the most emphasis being placed on ethnicity. Subsequent conference calls of the CCSC focused on discussions to address discrepancies in the community scores between reviewers.

After the CCSC selected 81 communities, the committee members conducted a second more informal review of the stratification variables (region, urbanicity, income, and proportion of ethnic minorities) with respect to the selected communities. Where the selected communities appeared to be inadequately distributed over these variables, CCSC members considered and discussed possible substitutions to enhance the distribution across strata while not significantly impacting the level of certainty in promising programs/policies that the CCSC had with the overall group. The CCSC was also asked to consider making substitutions that would enhance the coverage of states not already selected in the NBPS as long as again these substitutions did not significantly impact the level of certainty that the CCSC had with the overall group.

In most cases, the certainty community selected by the CCSC did not match the study definition of a community, i.e., it will cover a geographical area that contains multiple public high-school catchment areas. In these instances, we identified all public high-school catchment areas within the selected area and select one with probability proportional to size (where the measure of size

represents the 2010 U.S. Census population of children aged 3-15 years). These selection probabilities will be taken into consideration while developing the sampling weights at the community level, as described in Section B.1.b.4 below.

The CCSC members ultimately identified 81 certainty communities, with 5 alternates representing states that had not be included in the NPBS, for a total of 86 communities. Similar to the NPBS, we used a probability-based selection process to identify a single census tract within each of the 86 certainty community geographic areas – and then selected a public high school that was in closest proximity to the selected census tract to represent the certainty community.

***B.1.b.4 Description of Wave 2 Sample of Communities:***

As described in Sections B.1.b.2 and B.1.b.3, the community selection process yielded a National Probability-Based Sample that identified 195 unique census tracts, and a certainty community selection process that identified 86 additional census tracts. We also identified the public high school that was in closest proximity to those selected census tracts. There were several instances in which multiple census tracts in the sample were in closest proximity to the same public high school. This overlap was experienced both within the NPBS (selected census tracts from different strata were in proximity to the same public high school), and between the NPBS and the certainty community sample. Thus, the total number of communities (defined by the catchment area of a public high school) for the HCS was ultimately reduced from 281 (195+86) to 264 communities because of this overlap.

***B.1.b.5 Recruitment of School Districts and Procedures for Community Substitution in Wave 2***

School participation is vital to the success of the HCS. NORC and several study team members have a great deal of experience recruiting schools and students for research studies. The HCS recruitment process will follow a similar process as one undertaken by the Department of Education’s Early Childhood Longitudinal Study: Kindergarten Class of 1998-1999 (ECLS-K)<sup>3</sup>. First, letters will be to school district superintendents informing them of the HCS. Then research applications, if necessary, will be completed for each district. The applications will be submitted and telephone follow-up will be conducted. Similar to the ECLS-K, HCS has obtained the endorsement of national organizations representing school administrators, teachers, and parents. The national organizations are: The American Association of School Administrators, The National Education Association Health Information Network, The American Alliance for Health, Physical Education, Recreation, and Dance, and the National Parent Teacher Association. Once the school district approves participation in the HCS, contact will be made with individual schools. Based upon our early work with recruiting school districts into the HCS, we anticipate a 60-70% response rate for the school districts.

The goal within selected communities is to interview school personnel, identify potential study participants for in-home data collection and conduct environmental observations within 2 middle and 2 elementary schools (or their K-8 school equivalents) within each community, the minimal requirement is either 1 middle and 1 elementary school or 1 K-8 school. In the event that the

---

<sup>3</sup> National Center for Education Statistics. Early Childhood Longitudinal Study: Kindergarten Class of 1998-1999. ECLS-K base Year Public Use Data Files and Electronic Codebook. <http://nces.ed.gov/pubs2001/2001029rev.pdf>

HCS cannot meet the minimum requirement, the HCS will initiate a community substitution procedure to select a replacement community as follows:

1. If the HCS fails to gain access to schools at the district level – then all public high schools within that district will be eliminated from consideration in the HCS sample to avoid further non-response issues. Note that a considerable effort will be made to ensure that this is a rare event, and that the HCS has sequenced communities within the larger school districts associated with the 15 large-population certainty counties within the NPBS to occur later in the schedule to allow for a longer period to interact with these larger and potentially more complex districts.
2. If the community that requires substitution was selected from the NPBS from one of the strata (with no overlap with other strata or the certainty communities) – then another census tract will be selected probabilistically from the same strata and the public high-school in closest proximity will be selected as the replacement community.
3. If the community that requires substitution was selected from the NPBS from one of the 15 large certainty counties (with no overlap with the certainty communities) – then another census tract will be selected probabilistically from the same certainty county and the public high school in closest proximity will be selected as the replacement community.
4. If the community that requires substitution was selected from a certainty community (with no overlap with the NPBS) – then another census tract will be selected probabilistically from the same geographic area originally identified as the certainty community and the public high-school in closest proximity will be selected as the replacement community.
5. If the community that requires substitution was selected from the NPBS, but represents multiple strata from the NPBS (with no overlap with certainty communities) – then the multiple strata (groups of census tracts) will be combined and a new census tract will be selected probabilistically from the combined strata with the public high-school in closest proximity will be selected as the replacement community. This procedure will preserve the integrity of the NPBS.
6. If the community that requires substitution was selected from the NPBS, but represents overlap between the NPBS and the certainty communities – then we will first select a replacement community within the NPBS using the most appropriate approach cited above (either 2,3, or 5) in an effort to preserve the integrity of the NPBS. If resources permit and there is the possibility of selecting another community within the geographic areas identified as the certainty community – the HCS will also select a second replacement community following the approach cited in (4) above.

#### ***B.1.b.7 Temporal Ordering of Field Operations***

The HCS will schedule Wave 2 communities for field implementation (whether chosen as a certainty community or via probability-based selection) in a way that best facilitates operational efficiency and successful outcomes. For example, communities within large metropolitan areas such as Los Angeles, California or Chicago, Illinois will be visited slightly later in the schedule, allowing time for our community outreach team to work collaboratively with these larger, more complex school districts to gain entry into the selected schools. We have also utilized complex GIS analyses to identify clusters of communities that can be efficiently visited by the same data collection team – allowing these teams to perform data collection in multiple communities in a

serial sequence and the HCS to capitalize on the data collectors gaining experience with the study protocol over a longer period of time.

### ***B.1.c Recruitment and Selection of Schools and Child-Parent/Caregiver Participants***

The Wave 2 family recruitment approach will recruit schools to help identify families in the community who may be interested in participating in HCS. Using this strategy, the HCS will be able to directly associate the observations of the school physical and nutrition environments with the child and parent-level measures results since the sampled children will be attending the recruited schools.

#### ***B.1.c.1 Recruitment and Selection of Schools***

Upon receipt of District Office approval, schools will be contacted through mail or via email using a standardized invitation letter and enclosing a copy of the District Office approval for the study. Follow up calls will be made to each school to respond to any questions they may have and recruit them into the study. Participating schools must agree to: a) work with study staff to provide informational materials to families and distribute and collection interest forms to create the sampling pool for recruitment of families; and, b) allow HCS staff to conduct observations at the school. Written permission from the school principal will be required. School principals, or their designees, will also be requested to participate in a key informant interview. If a school refuses to participate in the study, or is non-responsive, then the next school on the list (of the same type – middle or elementary) will be contacted. A minimum of one elementary and one middle school, or one K-8 school, will be required within each community.

Schools will be selected using a probability-based sampling approach. For communities that were selected to represent a minority subpopulation (e.g., high proportion African American community or high proportion Hispanic/Latino community), Battelle will utilize a sampling method with probability proportional to size to select the elementary, middle or K-8 schools to approach for participant recruitment; where the measure of size is proportional to the number of students in the school that are either Hispanic/Latino or African American (depending on which strata the school represents). This will maximize the probability that schools with characteristics matching the strata will be selected. For all other communities, the measure of size for school selection will be the number of students.

We anticipate that for schools where the school districts have approved participation in the study, the response rate will be about 80%. The school response rate for public schools in ECLS-K was about 70%. The ECLS-K required a high level of school commitment which included in-school direct observations of children. In comparison the HCS will be much less burdensome to schools, thus we expect a higher response rate than that of the ECLS-K.

#### ***B. 1.c.2 Recruitment and Sampling of Child-Parent/Caregiver Participants from the Schools***

In schools that have agreed to participate in the HCS, children/parents will be recruited based on grade, gender, race/ethnicity, and the family's interest in participating. Each recruited school will receive a Toolkit with materials that can be shared with students' families to provide them with information on the study and assist in identifying those families with children in grades K-8 that may be interested in participating in the study. The Toolkit (see **SSA Attachment 4**) will be

available at least one month before any recruitment occurs in the school to allow time for these materials to be disseminated to families in advance of being contacted by the study team.

The Toolkit contains the following material:

- a. *HCS News Article*: A short news article in electronic format about the HCS that the school can include their PTA newsletter and/or other e-mail updates that routinely get sent, or are available, to parents of children in the schools.
- b. *HCS Brochures*: Brochures describing the study will be provided for students to take home to their parents/guardians. The brochures will explain the study's purpose, methods, and plans for implementation.
- c. *Parent/Guardian Letters and Interest Forms*: An invitation letter addressed to the child's parent/guardian and a study interest form will be sent home with children in their book bags to solicit parent's interest in learning more about the study. The interest forms will be printed in large easy to read typeface and have limited data items to be completed by the parent/guardian, including a name, contact number, and the grade, gender and race/ethnicity of the child attending the selected school. The student will be asked to bring the completed interest form back to school. This information will be used to develop the sampling pool from which we will recruit families for data collection. We will then employ stratified sampling methods within communities to prioritize the selection of minority families for participation in the HCS.

To assist with the dissemination and collection of the Toolkit materials, a school liaison from amongst the professional staff at each school will be identified. This individual will also serve as the "School Champion" for the study, helping raise awareness of the study in the school and community. To successfully function in this capacity, the selected individual must be highly regarded and trusted by both students and their parents. The prime responsibility for the liaison will be to serve as the point of contact for the study to parents/guardians and the students; this includes engaging students to take the informational materials and interest form to their homes and return the completed form back to the school, and responding to questions from parents/guardians on the study. The school champion/liaison will receive a \$50 gift to thank them for their assistance with the study. Each liaison will collect the returned interest forms from students and provide these to the HCS team.

***Sampling Child/Parent Participants within Households:*** The information collected from the interest forms will be compiled and used as the sampling pool for recruiting families. We expect that the response rate for families completing the interest form will be about 40%. Families who return the interest form will be called to screen for inclusion in the study. Of those children who return the interest form, we expect the response rate for recruitment will be 80%.

In order to help increase the return of interest forms children will be encouraged to return the interest forms whether or not they are interested in participating in the study. A small incentive (valued at no more than \$1 per student) will be provided to each student who returns their completed interest form. The HCS will work with the school liaison to best implement the process for distributing the incentives for returning completed interest forms.

A limitation of recruiting participants from those who respond to interest forms is that the findings will not be representative of student across the country and thus will not be generalizable to all U.S. public elementary and middle school children. The benefit of recruiting students through interest forms is that only those families who are interested in participating in the study will be screened thereby increasing efficiency of recruitment.

Households found to be outside the high school catchment area will not be considered as eligible for the study. Additionally, no more than one child participant who attends one of the recruited schools will be included in the selection process per household. Criteria for excluding children include:

1. no more than one child participant per household/family;
2. children who are institutionalized or are not ambulatory; and
3. children who have lived in the community for less than a year.

In each Wave 2 community, the study will recruit nine children within each grade level (K-8), totaling 81 children per community. The number of planned child and parent/caregiver home assessments for the HCS is detailed in Table B.1.2 below. The parents from whom we will collect anthropometric data are the biologic mother or father, or both; if the biologic parents are not available, anthropometric data will be collected from the caregiver.

**Table B.1.2 Number of Planned Child/Parent Assessments by Type of Community**

Year Of Wave 2 Data Collection	Type of Community Sample	Number of Communities	Participants Per Community			Participants Across Communities		
			Standard Protocol Only	Enhanced Protocol	Medical Records	Standard Protocol Only	Enhanced Protocol	Medical Records
Year 1  Year 3	Wave 2.1	104	72	9	57	7,488	936	5,928
	Wave 2.2	104	72	9	57	7,488	936	5,928
	Wave 2.3	56	72	9	57	4,032	504	3,192
<b>Total Number of Participants</b>						<b>19,008</b>	<b>2,376</b>	<b>15,048</b>

Table B.1.3 below shows the recruitment goal by grade and gender for communities not designated as high proportion minority communities (i.e., >30% African American or Hispanic/Latino residents), and Table B.1.4 shows the recruitment goal by grade, gender and race/ethnicity for communities designated as high proportion minority communities involved in the HCS. [Note that due to an odd number of children being recruited into the Wave 2 communities within each grade level (n=9), we created two sets of Wave 2 recruitment goals (Types A and B) in order to maintain gender balance across participants in these communities.]

**Table B.1.3 Recruitment Goals for Children by Grade and Gender for Communities not Designated as High-Proportion Minority Communities**

School Grade of Child at Enrollment	HCS Communities Wave 2 TYPE "A"		HCS Communities Wave 2 TYPE "B"	
	Females	Males	Females	Males
Grade K	4	5	5	4
Grade 1	5	4	4	5
Grade 2	4	5	5	4
Grade 3	5	4	4	5
Grade 4	4	5	5	4
Grade 5	5	4	4	5
Grade 6	4	5	5	4
Grade 7	5	4	4	5
Grade 8	4	5	5	4

**Table B.1.4 Recruitment Goals for Children by Grade and Gender for Communities Designated as High-Proportion Minority Communities**

School Grade of Child at Enrollment	HCS Communities Wave 2 TYPE "A"				HCS Communities Wave 2 TYPE "B"			
	Females		Males		Females		Males	
	Minority	Non-Minority	Minority	Non-Minority	Minority	Non-Minority	Minority	Non-Minority
Grade K	2	2	3	2	3	2	2	2
Grade 1	3	2	2	2	2	2	3	2
Grade 2	2	2	3	2	3	2	2	2
Grade 3	3	2	2	2	2	2	3	2
Grade 4	2	2	3	2	3	2	2	2
Grade 5	3	2	2	2	2	2	3	2
Grade 6	2	2	3	2	3	2	2	2
Grade 7	3	2	2	2	2	2	3	2
Grade 8	2	2	3	2	3	2	2	2

**B.1.c.2 Selection of Participants for Enhanced Protocol**

The HCS will include a random sample of study participants (one of nine, or approximately 11%) in Wave 2 that will participate in the Enhanced Protocol, which will involve more detailed nutritional and physical activity assessments.

**B.1.d Key Informants**

Community key informants, knowledgeable of programs and policies targeting childhood obesity in their communities, will be asked to provide information to document the evolution of these programs and policies. Approximately 40 potential key informants will be screened to identify 10 to 14 key informants within each community. The key informants will be selected such that they have a variety of roles and represent a wide array of interests within the community. Interviewing 10-14 key informants in each community will allow for a diverse set of perspectives and is cost-effective.

Key informants will include these types of individuals identified within several key settings/sectors:

- a) Schools (e.g., Principals of randomly selected Middle Schools/Elementary Schools, school health coordinator, parent-teacher organization);
- b) Health Organizations/Coalitions (e.g., chair of active community coalition, administrator/staff of local health department, administrator/staff of local hospital);
- c) Government (e.g., city manager/administrator, staff of parks and recreation, urban planner); and
- d) Non-Profit/Community Organizations/Service Agencies (e.g., administrator/staff of United Way, organization with knowledge of child care organizations, youth organization, community foundation, neighborhood organization, Chamber of Commerce, faith organization).

Table B.1.5 below provides a listing of the planned priority (and alternate) for selection of key informants by setting.

**Table B.1.5 Priority/Alternate Community Key Informants Targeted for the HCS**

SETTING/SECTOR	PRIORITY/ALTERNATE KEY INFORMANTS
SCHOOLS	PRIORITY: Principal of up to 2 randomly selected Middle School
	PRIORITY: Principal of 2 randomly selected Elementary Schools
	ALTERNATE: School health coordinator
	ALTERNATE: PE coordinator
	ALTERNATE: Food service coordinator
	ALTERNATE: Additional principal of middle school or principal of high school
HEALTH ORGANIZATIONS/ COALITIONS	PRIORITY: Chair of active community coalition (if present)
	PRIORITY: Administrator/staff of local health department
	PRIORITY: Administrator/staff of local hospital
GOVERNMENT	PRIORITY: City manager/administrator OR Urban Planner
	PRIORITY: Head/ staff of parks and recreation
	PRIORITY: Administrator/staff of United Way
NON-PROFIT/ COMMUNITY ORGANIZATIONS/ SERVICE AGENCIES	PRIORITY: Director/staff of community foundation with health mission
	PRIORITY: Director/staff of organization with knowledge of child care organizations (if data not available from other sources)
	ALTERNATE: Director/staff of child or youth organization
	ALTERNATE: Leader of active neighborhood organization
	ALTERNATE: Head of Chamber of Commerce
	ALTERNATE: Leader of active faith community

The first potential key informants in a community will be identified using the following approaches:

- 1) Utilize the Market Data Research (MDR) database to identify contact information for personnel associated with public schools within the catchment area and corresponding district,

- 2) Web-based searches to identify people within the municipal government (e.g., a representative from parks and recreation, city/county health department), and other organizations within the community (e.g., local YMCA, Chamber of Commerce, United Way, local hospital, etc.) to consider as possible key informant to interview. The initial focus of these searches will be to identify key informants in each sector/setting as listed in Table B.1.5; and.
- 3) Review of information gathered as part of the Certainty Community identification process (e.g., Program specific information that can be used to identify appropriate affiliated key informants).

From these multiple sources, prospective key informants and as much contact information as can be gathered will be entered into the Information Management System (IMS). Priority key informants will be identified in the database to facilitate the review of the list and a Decision Tree will be developed delineating the type(s) of key informants to try and reach should the primary key informant listed as priority be unavailable or unwilling to participate. This will help inform the community liaison's outreach to key informants across a broader range of organizations/roles within a community and within the "priority" sectors/roles.

All prospective key informants will be screened by telephone to: describe the study, determine the key informant's eligibility, and invite the individual to participate in the study as a key informant. If requested by the potential key informant during these screening calls, an informational letter (see **SSA Attachment 10**) and a study brochure (**SSA Attachment 11**) will be mailed to provide information on the study and the role of a key informant. (Please see **Attachment 12** for the screening script and **SSA Attachment 13** for the recruitment script).

Once a prospective key informant agrees to participate in the study, an appointment for an interview will be scheduled to take place either in-person or over the phone at a time convenient for the key informant and documentation on community programs/will be requested. Each key informant interview is planned to take no more than one hour and thirty minutes (please see **SSA Attachment 13** for the interview questionnaire).

During the screening call the community liaison will also ask the prospective key informant to identify additional appropriate key informants from other sectors to contact; this snowball sampling method will assist the community liaison in identifying key informants from a variety of sectors to obtain a more comprehensive overview of the community's programs and policies. Additionally, community liaisons will keep track of which types of key informants in each community they have contacted/interviewed and use this as a visual cue to help ensure broader coverage of key informants across multiple sectors. This information will be used in assessing the completeness of data collection across the sectors which recruited key informants represent.

At the time of the interview, the Battelle community liaison will explain the study, review the consent document, and answer any questions the key informant may have. Following this, the key informant will be asked to sign the informed consent form prior to the initiation of the structured interview. When the respondent is unable to participate in an in-person interview and completes a telephone interview instead, a verbal consent script will be read to the key informant by a Battelle community liaison before the interview begins. Battelle community liaisons will be

trained in Human Subjects Research; therefore, they will know the guidelines regarding what qualifies as informed consent.

### ***B.1.e Power Calculations***

The inclusion of 264 communities in Wave 2 of the study will allow the HCS to observe a considerable degree of natural variation in approaches and intensity of effort in childhood obesity programs and policies. The involvement of 81 children in the K-8 grade range within each of the communities will provide the HCS with adequate power to observe associations between community-level efforts and childhood obesity measures.

Researchers may be interested in the association between different program/policy components and childhood obesity measures for all ages combined and within a particular age group using a variety of approaches. One approach would use all data measured when a subject's age was within the age group of interest regardless of the subject's age at the baseline assessment. For example, if a participant who was 15 years old at the baseline assessment was found to have medical records containing BMI information when he/she was seven years old, those records could be included in an analysis focusing on the 7-8 year-old age group. A second approach would use all data ever measured on an individual whose baseline age fell within the age group of interest, enabling specific cohort analyses. For example, a study focusing on 7-8 year-olds would include all data collected on those children that were ages 7 or 8 at the baseline assessment. Our simulation study provides insight into the power that the HCS may have while employing these age-specific analysis approaches.

Table B.1.6 lists the estimated difference in BMI as a percentage change that might be detected for a one-unit difference in intensity score (i.e., from 0 to 1, that is, from having no intensity to the maximum intensity measured), assuming a cross-sectional analysis with 80 percent power. The simulations cover two scenarios in which only a small fraction of communities will have implemented programs/policies with high intensity, as well those in which program intensity varies uniformly across communities. Our analyses indicate that a cross-sectional analysis with 264 Wave 2 communities and about 81 children per community, spread evenly across grades K-8 might have adequate power to identify programs/policies that reduce BMI by approximately 3.7-5.4% across all age groups, and 5-8% in specific age groups, if certain assumptions about the underlying participants are met. For example, a community with no programs/policies (score = 0) has an average BMI of 23.0, while a community with the most intense programs/policies (score = 1) has an average BMI of 21.767, resulting in a true difference of 5.36%. The HCS will detect that 5.36% difference as statistically significant 80% of the time, assuming that only a small fraction of communities will have implemented programs/policies with high intensity (second scenario). That is, the HCS will conclude it is an actual difference rather than noise given the sample of children selected. Additional details on the methods used to estimate power can be found in **Attachment 3**.

**Table B.1.6 Estimated % Difference in BMI that Can Be Detected for a One-unit Difference in Intensity Score Assuming a Cross-sectional Analysis with 80 % Power**

Age group for Analysis	% Difference in BMI that can be detected for a one-unit difference in intensity score	
	Uniform distribution of program intensity across communities	Few communities implement at full intensity, most communities at low intensity
	beta(1, 1) Uniform	beta(0.5, 2) Skewed
All Ages	3.69%	5.36%
Ages 4-6	4.66%	6.79%
Ages 7-8	4.66%	6.63%
Ages 9-10	4.81%	6.89%
Ages 11-12	4.69%	6.72%
Ages 13-14	5.70%	7.90%
Ages 11-14	4.45%	6.35%

Table B.1.7 provides the estimated power that the HCS will have to detect changes in BMI associated with a one-unit change (i.e., change in intensity score from 0 to 1) in program intensity, assuming a longitudinal analysis that integrates the historical height and weight data obtained via medical record abstraction. These designs have higher power than our cross-sectional designs, because they add repeated measures for the study participants. These longitudinal models allow the HCS to investigate how childhood obesity measures are associated with changes in community-based programs and policies over time. The longitudinal analyses that combine all available BMI data across all ages are able to detect smaller effect sizes due to the large number of children in the study, and the projected availability of repeated measures BMI data from approximately 70% of participants who will provide historical BMI data through medical record review.

**Table B.1.7 Estimated Power to Detect a 3% Change in BMI for a One-unit Difference in Intensity Score Assuming a Longitudinal Analysis**

% Change in BMI	Age group for Analysis	Power (Probability that detected change in BMI is significant $p \leq 0.05$ )	
		Uniform distribution of program intensity across communities	Few communities implement at full intensity, most communities at low intensity
		beta(1, 1) Uniform	beta(0.5, 2) Skewed
0.7% Change	All Ages	91%	84%
<b>Age-specific analyses using age at time of BMI recorded in medical chart or in-person visit</b>			
3% Change	Ages 4-6	>95%	>95%
	Ages 7-8	>95%	>95%
	Ages 9-10	>95%	>95%
	Ages 11-12	93%	83%

% Change in BMI	Age group for Analysis	Power (Probability that detected change in BMI is significant $p \leq 0.05$ )	
		Uniform distribution of program intensity across communities	Few communities implement at full intensity, most communities at low intensity
		beta(1, 1) Uniform	beta(0.5, 2) Skewed
	Ages 13-14	69%	48%
	Ages 11-14	97%	92%
<b>Age-specific analyses using age at in-person visit</b>			
1.5% Change	Ages 4-6	>95%	94%
	Ages 7-8	74%	66%
	Ages 9-10	66%	57%
	Ages 11-12	82%	67%
	Ages 13-14	82%	70%
	Ages 11-14	93%	83%

We also examined the power to detect the associations between community programs/policies and binary measures of physical activity or nutrition. Table B.1.8 lists the percentage point difference in the probability that a participant has the binary outcome that could be detected for a one-unit difference in program intensity, assuming a cross-sectional analysis. The results provided in Table B.1.8 demonstrate that the study will have 80% power to detect differences of 4-5 percentage points associated with a one-unit difference in community program/policy intensity score across all age-groups combined, and differences of 7-14 percentage points across most specific age groups. For example, suppose 50% of the children in a community with no programs/policies (score = 0) consumed a sugar-sweetened beverage in the last 7 days, while 44.89% of children in a community with the most intense programs/policies (score = 1) consumed a sugar-sweetened beverage in the last 7 days. This is a true difference of 5.11 percentage points. The HCS will detect that 5.11 percentage point difference as statistically significant 80% of the time, assuming that only a small fraction of communities will have implemented programs/policies with high intensity (second scenario). That is, we will conclude it is an actual difference rather than noise given the sample of children selected.

**Table B.1.8 Estimated Percent Difference for Binary Diet or Physical Activity Measure for a One-unit Difference in Intensity Score Assuming a Cross-sectional Analysis with 80 % Power**

Age group for Analysis	% Difference for a Binary Diet or Physical Activity Measure that can be Detected)	
	Uniform distribution of program intensity across communities	Few communities implement at full intensity, most communities at low intensity
	beta(1, 1) Uniform	beta(0.5, 2) Skewed
All Ages	3.98%	5.11%
Ages 4-6	8.48%	11.05%
Ages 7-8	8.30%	10.86%

Age group for Analysis	% Difference for a Binary Diet or Physical Activity Measure that can be Detected)	
	Uniform distribution of program intensity across communities	Few communities implement at full intensity, most communities at low intensity
	beta(1, 1) Uniform	beta(0.5, 2) Skewed
Ages 9-10	8.69%	10.49%
Ages 11-12	7.94%	10.40%
Ages 13-14	10.44%	14.10%
Ages 11-14	6.58%	8.26%

Based on a number of assumptions, power for this study was estimated using Monte Carlo simulations in R (version 2.14.2). Multiple datasets were simulated and analyzed using the mixed model routines in the lme4 package. The point estimate of the program effectiveness parameter was recorded after analyzing each simulated dataset. The standard deviation of the point estimates over all datasets was used to approximate the standard error. We related power to effect size using the following equation:

$$\beta_1 = SE \times (z_p + z_{1-\alpha/2})$$

where  $SE$  is the standard error approximated from the simulations,  $P$  is the power,  $\alpha$  is the significance level of the test, and  $z_u$  is the  $100 \cdot u^{\text{th}}$  percentile from a standard normal distribution.

We will use the Enhanced Protocol measures, which yield more detailed data, and apply statistical modeling techniques to develop valid statistical inferences on the relationship between the more complex and burdensome nutritional and physical activity measures and measures of community program and policy intensity. We will apply what we learn from the Enhanced measures to the Standard measures, being able to statistically adjust for bias and error in these less complex and less detailed measures. The two-stage statistical analysis being implemented will thereby improve the study's power without increasing burden for all child and parent/caregiver respondents.

Several assumptions were made in estimating power and detectable differences. In the cross-sectional power studies, we utilized only the BMI data collected during the home visit. There are no data available to examine prospective changes over time. For the retrospective, longitudinal study design, we expect 70 percent of the children to have medical records available that contain height and weight information. The number and timing of the previous medical records were simulated using abstracted medical record data from Wave 1 participants (OMB Notice of Approval 0649 dated 1/30/2012). The simulations assumed that the average number of records per child was 5.66 (standard deviation = 5, minimum = 1, maximum = 23); that the average time since the earliest record was 6.07 years (standard deviation = 3.95, minimum = 0.47, maximum = 15.01); that the average time since the most recent record was 1.97 years (standard deviation = 2.11, minimum = -0.14, maximum = 8.82); and that the average spacing between records was 1.19 years (standard deviation = 1.18, minimum = 0.08, maximum = 6.08).

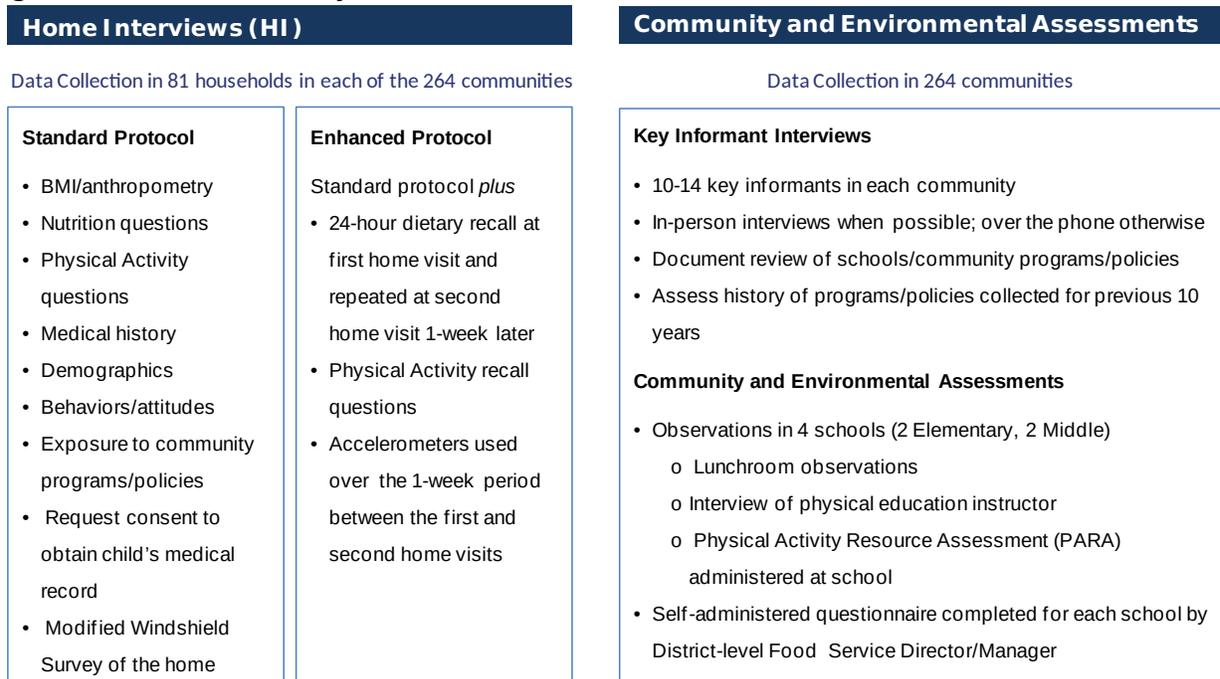
Each child was assumed to receive an assessment of the binary measure of physical activity or nutrition. This would be the only data point available for this type of response.

For the longitudinal analyses, we developed a model to simulate changes in program/policy intensity over time. Program/policy intensity was assumed to follow a logistic curve with a slope that could be positive (allowing program intensity to increase over time) or negative (allowing program/policy intensity to decrease over time). Each community’s slope was randomly sampled from a normal distribution with a mean of zero and a standard deviation of 1.25.

## B.2 Procedures for the Collection of Information

Data collection within the selected communities will be hierarchical by nature, and will include two primary components aimed at (1) documenting how childhood obesity programs and policies have evolved within the community over time dating back ten years, and (2) assessing current anthropometrics (height, weight, BMI), physical activity, dietary behaviors, a longitudinal history of BMI abstracted from medical records, and a variety of other important covariates and confounders on a sample of children from within the community. This data collection will occur over a 1.5-year period. Figure B.2.1 provides an overview of data collection activities, with additional detail provided in the supporting documentation.

**Figure B.2.1 HCS Primary Data Collection Activities in Wave 2, 2013-2015**



### B.2.a Community-Level Data Collection

Data collection at the community level will include a retrospective review of programs and policies to document how childhood obesity programs and policies have evolved within the

community over the previous ten years, and also will characterize the current state of these programs. Documentation of the evolution of programs and policies within communities will occur via key informant interviews using multiple modes of data collection (telephone interviews, web-based questionnaires, in-person meetings). In addition, these interviews will be supplemented by document retrieval and abstraction by research staff. Please refer to **SSA Attachment 13** for the key informant interview protocol.

Data collection at the community level will also include assessments of the broader community, as described in Section B.2.a.3 below.

### ***B.2.a.1 Document Review***

One component of this data collection will be the initial gathering and comprehensive review and documentation of program and policy-related information. During the key informant interviews in all communities, a Battelle community liaison – a full-time, trained research staff member - will request to review documents developed by or about relevant community programs and policies. The types of documents we will review include: publicly available legislative hearing documents; annual program reports; management information system (MIS) reports often used by funders; available and relevant Requests for Proposals (RFPs) and responses to RFPs that may have been prepared within the time frame of interest; agency-wide reports such as Healthy People 2000 and 2010; coalition reports from relevant organizations, schools, and other institutional and community wellness policies.

Using a field-tested document abstraction instrument, data will be reviewed and then entered into an electronic form.

***B.2.a.1.a. Review of Online Databases:*** We will also review online databases from various sources prior to initiation of key informant interviews in each community to supplement the document review<sup>4</sup>. Preliminary testing of a protocol for conducting document abstraction using online databases reveals that there are two categories of databases that should be reviewed that may serve as a prompt during the key informant interview: databases that contain descriptions of community programs/ policies (four databases) and databases that indicate participation in national initiatives (five databases). The following priority databases will be reviewed:

- *Category 1: Databases to be reviewed for descriptions of community programs/ policies*
  - o *Community Commons*, a free-use, project-level data visualization and geographic information system (GIS) mapping platform used to track population health program/policy adoption at the site for implementation.
  - o *Healthy Counties Database*, a national, county-level database containing enacted programs and policies intended to support overall community health.
  - o *National Conference of State Legislatures (NCSL) Bill Summaries Database*, a comprehensive database featuring all state-level legislation, discriminated by topic, including nutrition and physical activity.

---

<sup>4</sup> We will conduct this review of online databases for the first 12 communities visited in Wave 2; review of online databases for the remaining Wave 2 communities will be conducted if additional funding is obtained through a diversity supplement.

- o *Alliance for a Healthier Generation*, this is an initiative that addresses childhood obesity nationwide.
- *Category 2: Databases to be reviewed for participation in national initiatives*
  - o *Safe Routes to School*, this website has information nation-wide that include initiatives for getting to school safely.
  - o *Healthy Kids, Healthy Communities*, this national initiative supports healthy food and physical activity programs in 50 different communities.
  - o *Active Living by Design*, a nationwide initiative that promotes active living and healthy eating.
  - o *Community Transformation Grants*, this website contains a list of awardees that received CDC grant awards in implementation and capacity building.
  - o *Communities Putting Prevention to Work*, is a CDC based initiative that supports 50 communities in addressing obesity and tobacco use.

Each online database in Category 1 will be reviewed in the same manner as described in Section E.3.1 and related documents. Databases in Category 2 may be reviewed in abridged manner to discern if the databases list the participation of any participating community in the initiative. This can be used as a possible prompt during key informant interviews (e.g, “ Review of online resources indicated that Greenville Elementary has a Safe Routes to School program, can you tell me more about this?”).

### ***B.2.a.2 Key Informant Interviews***

In each community, the Battelle community liaison will identify and conduct interviews with key informants associated with and/or knowledgeable of relevant community programs and policies. We will use standardized data collection instruments to gather characteristics for each program and policy operating in the selected communities, such as: target population, including age, sex, location; target focus area; who delivered the program/services; level of funding and how it was funded; date initiated and ended, when applicable; number reached, and who was reached (i.e., providers trained and recipients of service); nature of collaborations; and any outcomes examined.

Key informants will answer a similar set of interview questions for each specific local program and/or policy identified during key informant interviews. Data will be combined from different key informant interviews for the same program or policy.

### ***B.2.a.3 Other Community Assessments***

Community assessments will include interviews with additional school administrators/personnel, community key informants, participant perceptions of the school and home environments, GIS data, and direct observations of schools to collect program/policy and environmental data. The following assessments will be conducted during the in-person visit in all 264 communities.

***Nutritional Environment:*** Observational assessments of the nutritional environment will be conducted in all the schools recruited in each community. The Battelle community liaison will observe the school’s lunch period and complete an observation form during the school visit; additionally, the Food Service Administrator/Manager at the district level will be requested to complete a brief web-based survey for each recruited school in their district to capture

information on the school lunch program. A brief, web-based survey that requests information about wellness policy implementation at the school will be provided to the school liaison during the school recruitment process to complete and return to Battelle. In addition, questions will be asked of the principals of the recruited schools during the key informant interview, should they agree to participate.

The surveys and observations will provide information regarding:

- school characteristics -- student enrollment, meal program participation, free and reduced price meal eligibility, open campus status, participation in key federal school nutrition and food programs, degree and extent of scratch food preparation, extent of wellness policy implementation and how long in place, if and when changes have been made to the school meals, meal facilities and competitive foods;
- characteristics of reimbursable lunch options -- number and type of entrees, beverages, whole grain vs. refined grain products, fruits, vegetables, dessert, and snack foods;
- characteristics of dining facilities -- availability, ambiance, size;
- cafeteria staff interactions with students;
- length of the lunch period and average length of time to obtain lunch;
- recess during lunch time;
- water availability; and
- competitive foods -- what, where, and how much is offered.

These data will be used to characterize the past ten years of the school food environment and validate and add detail to stakeholder reports of school-based efforts. Furthermore, the data will be used for cross-sectional and retrospective, analyses of associations with anthropometric and behavioral measures.

***Physical Activity Environment:*** Observational assessments of the physical activity environment will be conducted in the recruited elementary and middle schools in each community. The Battelle community liaison will interview a member of the physical education staff at the school, and will also observe the school's physical activity resources using the Physical Activity Resource Assessment (School PARA) form. A brief, web-based survey that asks about physical activity policies and practices within the school including physical activity and punishment practices, collaborations with community partners, and walk-to-school practices will be provided to the school liaison during the school recruitment process to complete and return to Battelle. In addition, questions will be asked of the principals of the recruited schools during the key informant interview, should they agree to participate.

Furthermore, key informants who are particularly knowledgeable about parks and recreation and other physical activity resources in the community will answer additional questions. As part of the key informant interview, the Battelle community liaison will ask a few supplemental interview questions to gather information about community resource availability and accessibility, physical activity related community collaborations, park and trail use and general features, and other information related to physical activity resources in the community.

The physical education interview will gather data about physical activity resources and facilities on school campuses, the provision of physical education, recess, and other physical activity

opportunities at schools, the community partnerships established for providing physical activity opportunities, and the norms and culture for physical activity at the schools.

The School PARA will be conducted for outdoor features of the environment related to physical activity. This form characterizes the features, amenities, and incivilities of the physical activity environment(s) in the school, the hours, availability, and capacity of the facilities, and size and cost of use of these environments.

## ***B.2.b Data Collected at the Child/Parent Participant Level***

### ***B.2.b.1 Anthropometrics***

Anthropometric measurements will be taken at the household visit for all participants. Anthropometric measurements to be collected include the height, weight, and waist circumference of the child and the height and weight of the biological parent(s) either via measurement (if available during home visit), or self- or proxy-report. Measurements will be made according to the NHANES protocol, recorded in metric units (centimeters and kilograms), and measured to the nearest .1 cm and .1 kg. BMI will be calculated by dividing weight (in kilograms) by the square of height (in meters) [kg/ m<sup>2</sup>]. Height and weight will be measured twice during each home visit. For an 11% randomly selected sample of participants who receive the Enhanced Protocol, measurements will be conducted at the first home visits and again approximately one week later for quality control purposes as a measure of inter- and intra-observer reliability.

### ***B.2.b.2 Medical Record Review***

Medical records for approximately 70% of the child participant will be abstracted to develop longitudinal BMI trajectories for up to the previous 10 years. Any indication of nutritional, physical activity, or sedentary activity counseling will also be abstracted from medical records. The presence of other chronic conditions and prescribed medications for those conditions (e.g. asthma, diabetes) will also be abstracted. At the conclusion of each community assessment, our subcontractor, Examination Management Services Inc. (EMSI), will contact one medical provider for each selected participant whose parent/guardian provided consent to access medical records. Providers and participants will be selected based on a two-step algorithm. First, a single provider for each participant is chosen; if multiple providers are listed on the medical record release form, the provider most likely to give the best information is determined by calculating the expected number of height/weight measurements taken by each provider *while the participant has lived in the community* and comparing these among the providers. Second, participants (and their corresponding providers from step one) are selected for medical record review by computing a score based on the expected number of measurements taken while living inside and outside the community, the race/ethnicity of the child, whether the child was selected for the Enhanced Protocol, and income; a higher score corresponds to a higher likelihood of being selected for medical record review (see **SSA Attachment 20** for details regarding the selection of medical providers and participants). Approximately 70% of all child participants will be selected, with the goal of retrieving medical records for all of the selected children; if records cannot be retrieved for some of these child participants, consideration will be given to selecting alternate child participants so that medical records can be retrieved and abstracted for close to the target of 70% of all child participants. The EMSI data abstractor will load the

information into an electronic form to be uploaded to the secure study data repository at the conclusion of the abstraction process.

### ***B.2.b.3 Nutritional Assessments***

***Standard Protocol:*** Information on food and beverage intake, food patterns and behaviors, breastfeeding, household food insecurity, and perceived social support, perceived home environment, and perceived community environment regarding healthy eating will be collected on children enrolled in the study. For children six years or older, questions will also be asked about the perceived school environment regarding healthy eating. For children 12 years and older, additional questions will be asked regarding weight-based teasing and meal skipping for weight control. Questions will be age appropriate, and either be self-administered (child and parent/caregiver questions) or parent/caregiver assisted, depending on the age of the child.

***Enhanced Protocol:*** In addition to the standard nutritional assessments, participants selected for the Enhanced Protocol will complete two 24-hour dietary recalls 8 to 10 days apart using the ASA24-Kids (Automated Self-Administered 24-hour Recall-Kids). This online web-based 24-hour dietary recall is based upon NCI's ASA24 and has been adapted for use with children, containing considerably fewer probes. The ASA24 is a modified version of the U. S. Department of Agriculture (USDA)'s Automated Multiple Pass Method (AMPM), which utilizes multi-level food probes to obtain estimates of types and amounts of food consumed on the day prior to the visit. The ASA 24-Kids has been developed for self-administration, and will be completed at the first and second home visits for Enhanced Protocol participants. The field interviewer will log on and enter the child's ID, note the date and time the interview commences, and then turn over the computer to the primary respondent. The primary respondent, along with the secondary respondent (when applicable), will use the computer to enter the information prompted by the online mascot throughout the interview. The field interviewer will be trained to give a neutral introduction and clear instructions to the parent/caregiver and child regarding who is to respond and to encourage interchange to obtain the most accurate information about the child's food intake on the previous day. The field interviewer will sit with the primary and secondary respondent and assist with any difficulties experienced in searching for or finding foods, remind respondents to consider commonly forgotten foods and beverages as well as those consumed outside of the home, and encourage the respondent to complete all aspects of the interview in a timely fashion.

Data from the 24-hour dietary recall will provide estimates of total intake for all nutrients available in the Food and Nutrient Database for Dietary Studies (FNDDS), food groups included in the Food Equivalents Database (MPED) and selected other food sub-groups such as sugar-sweetened beverages. The purpose of collecting two recalls on each child is to enable adjustment for within person variability. The adjusted food and nutrient estimates from the ASA24-Kids will be used to calibrate estimates on children in the entire sample obtained from the Household Interview dietary screener.

### ***B.2.b.4 Physical Activity and Sedentary Behavior***

**Standard Protocol:** All participants will answer questions related to their physical activity (e.g., types of activities; including intensity, frequency, and duration) at home, at school, and in the community, *during the previous week*, and their parents will be asked questions related to their children's activities as well as physical activity resources available at home and in their community. Questions will be gender specific and age appropriate, and either be self-administered (child and parent/caregiver questions) or parent/caregiver assisted, depending on the age of the child.

**Enhanced Protocol:** In addition to the Standard Protocol, the child participants will be asked to recall and describe the activities they participated in at home, at school, and in the community, *the previous day* using a similar structured instrument based on gender and age. Participants will also be given an accelerometer at their first home visit. Accelerometers provide an objective measure of physical activity by detecting movement and intensity of activity. The device will be worn by participants except while bathing, swimming, or sleeping, for the week in between home visits.

#### ***B.2.b.5 Demographics, Family Medical History, Direct Observations***

All parents/caregivers will answer the following demographic questions related to the child and parent/caregiver: age, race, ethnicity, marital status, country of origin, education, language, employment, and family income. In addition, the child's and parent/caregiver's height and weight will be measured (or reported if adult refuses measurement).

A series of questions are included in the Standard Protocol, aimed at assessing the medical history for the child participant related to:

1. Issues related to the participant child's access to healthcare;
2. Medical conditions that may alter diet for the participant child (diabetes, celiac disease, anorexia, bulimia, etc.);
3. Medical conditions that alter physical activity for the participant child (disabilities, recent accidents (e.g., broken ankle), etc.);
4. Medical conditions that alter ability for the participant to self-complete aspects of protocol (cognitive deficits, Down's syndrome, dyslexia, etc.).

#### ***B.2.b.6 Direct Observations of the Child's/Family's Neighborhood***

In all communities, direct observations of the child's home will be completed by study staff.

The field interviewers will complete a five-item modified windshield survey during the baseline visit. EMSI field interviewers will rate features of the social and physical environment on the street segment associated with each child's home address. A street segment is defined as the street in front of the home, from intersection to intersection, not to exceed 0.5 miles. In the instances where the street segment exceeds 0.5 miles, the field interviewers will be instructed to consider the street segment that is contained within 0.25 miles from the home in either direction or to the nearest intersection, whichever is closer.

A paper form of the modified five-item windshield survey will be completed when the field interviewer arrives at the child's home and entered into the study database following the home visit. The modified five-item survey is expected to take less than five minutes to complete. GPS coordinates will also be obtained for each participant's household to match to GIS data.

### **B.3 Methods for Maximizing Response Rates**

As mentioned earlier, we are using a school recruitment technique similar to the ones used by ECLS-K. Once schools have agreed to participate in the study, we will recruit students via interest forms distributed in schools. All data regarding contact success, refusal to screen and/or enumerate household members, refusal to participate for one or more eligible children (including reasons for non-participation) will be maintained to calculate response rates for the study. Response rates will be calculated as the fraction of eligible participants who elect to complete the data collection protocol, and will be calculated for child/parent participants as well as key informants.

Every effort will be made to maximize the response rates for key informant interviews, parent and child surveys, as well as acquisition of medical records by the participating child's medical provider. Based on our review of the current literature as well as our experience in the field with prior studies, the approaches that we will use to maximize the response rates, are described below.

**Key Informant Interviews:** Community key informants include, for example, elementary and middle school principals, directors of programs involved with children health issues, and leaders of local government programs and non-profit community organizations. It is anticipated that participation will be relatively high among this group given their a priori interest in helping improve the health and well-being of those living in their community. Furthermore, as this group will be comprised through a snowball sampling technique, we expect that the referral by one community leader to another will improve succeeding participation rates. To maximize response rates among this group, upon scheduling the initial appointment, we will maintain contact with a confirmation letter, including printed information about our study (see **SSA Attachment 10** and **SSA Attachment 11**), and follow-up confirmation telephone calls. Our protocol calls for conducting interviews with key informants in person, which is likely to improve our response rate, although when necessary due to scheduling or other conflicts, we will complete interviews over the telephone. All interviews will be conducted by the Battelle community liaison – a qualified and well-trained interviewer, who is well versed on the community and its programs/policies, and who is comfortable speaking with high-level community leaders and able to adeptly answer any study related questions. As we will also be requesting specific program/policy documentation from this group, we will follow-up on this request through our confirmation letter, subsequent telephone calls, and during our in-person interview meeting as needed.

**Parent/Child Surveys:** At each school that agrees to participate in the study a school liaison will be selected from amongst the professional staff to serve as a “Champion” for the HCS study. The prime responsibility for this champion will be to serve as the point of contact for the study to parents/guardians and the students, and be responsible for the interest forms. In all communities, the school liaison will serve a key role in supporting the recruitment of families into the study and will therefore need to understand the diverse needs of the families and the community; this

will be especially critical in the high proportion minority communities, particularly the Hispanic/Latino Communities. We will therefore request that the school principals in these communities select a Champion who understands the cultural and social dynamics of their particular community, has connections within and is trusted by this community, and can help in our recruitment efforts in and outside of the school environment by raising awareness of the study through their existing networks. The selected individual should also be someone whom the students and their parents respect, to whom they relate well, and someone they feel is easily accessible to respond to their questions or concerns regarding the study. The use of a School Champion should help raise awareness and trust of the study among potential study participants, thus encouraging participation and overall response rate.

The School Liaison/Champion will also be provided with a Recruitment Toolkit containing informational materials as well as family-specific letters, brochures and sign-up forms for parents/guardians to express an interest in participating in the study. These materials have been developed and carefully reviewed to ensure ease of readability, cultural sensitivity, and low burden for the potential participants. The informational materials will emphasize the study's role in assessing programs and policies that aim to improve the health of children, and the benefits children will receive by participating in the study. The information obtained through returned interest forms will be used to develop the sampling pool from which we will recruit families for data collection. This self-selection process will improve the likelihood of successfully recruiting families and completing the home visit.

Children will be encouraged to return the interest forms whether or not they are interested in participating in the study. A small incentive (valued at no more than \$1 per student) will be provided to each student who returns their completed interest form. The HCS will work with the school liaison to best implement the process for distributing the incentives for returning completed interest forms.

Completion of parent and child surveys in the home administered by well-trained field interviewers, rather than having the parents respond to a survey mailed to their home, is likely to result in a high response rate. In addition, when necessary, interviews (as well as screening/recruitment calls) will be available to be conducted in Spanish. Similar to the NHANES protocol, should neither English nor Spanish be spoken, the household member completing the family recruitment process will be asked to identify a family member or neighbor, aged 18 or over, who can be present during the home visit to translate to the interview questions into the necessary language. This will help maximize our initial enrollment and follow-up response rates. Furthermore, the offering of an incentive, though small, to both participating parent and child, is anticipated to further improve the response rate.

**Medical Provider Records:** It is anticipated that the number of records to be requested of any one provider will be relatively few. In addition, because we have budgeted to reimburse any administrative fees the providers may charge for copying or providing these records, and because our partner, EMSI, has a great deal of experience obtaining patient medical records from providers, we anticipate our response rate to this particular data collection effort to be relatively high.

#### **B.4 Tests of Procedures or Methods to be Undertaken**

All data collection instruments have been tested in house, or where possible, in the field, by members of the research team to ensure that timing will align with the estimate of respondent burden. While no further refinements are anticipated to the study design and protocol, if any further refinements to sampling and selection procedures, field protocols, or data collection instruments are made, we will inform OMB of all changes and provide updated and final documentation and instruments.

#### **B.5 Training and QA/QC Methods**

Experienced trainers from Battelle and its partners will oversee and/or conduct all staff training, including training of the Battelle community liaisons on conducting face-to-face interviews, document abstraction and community observations, and the field interviewers on conducting the home visits and accurately use and record measurements from the various anthropometric measurement devices (e.g., scale, accelerometer). The Battelle community liaison and independent QA/QC staff will also be trained on how to conduct quality control of the field interviewers once they begin the home visits.

Training sessions will be conducted using a number of modalities, including in-person one-to-one and group sessions (particularly for the training on how to conduct face-to-face interviews with key informants), on-line individual coursework, video viewing, and live webinars. As part of our training and quality assurance protocol, all study staff will also receive a personal training and reference manual, and will complete human research ethics trainings and confidentiality trainings.

#### **B.6 Consultations and the Project Team**

Per OMB's suggestion the HCS project team consulted with Chris Chapman, a survey sampling statistician at the National Center for Education Statistics, on the study design. Mr. Chapman reviewed the HCS study design and felt that the study questions could be answered with the current design. However, he indicated that the sample of students was not nationally representation and that the findings could not be generalized to U.S. public school elementary and middle school children. Below is a summary from Mr. Chapman regarding his thoughts on the HCS design:

“Studies simultaneously relating a wide range of policy initiatives to children’s weight have not been undertaken and would be very useful from a research and policy perspective. Childhood obesity is a pressing problem in this country, and there have been a large number of initiatives put into place to help address the issue. In some cases, the initiatives are new and specifically targeted to reducing obesity (e.g., shifting school lunch program offerings), in other cases policies have been in place for many years but are being adjusted to focus more intensively on helping avoid obesity in children (e.g., public park and recreation programs). Understanding how a program relates to childhood obesity in isolation and in tandem with other potentially overlapping programs would help set the stage for more targeted and in depth studies of policies that show the most promise. That is, a broad ranging study such as this one would be well positioned to inform considered, clear, and targeted interventions to help prevent or reduce our long standing weight problem. Without the more extensive environmental study, researchers and policy makers

will not have the broader perspective of how eclectic sets of local initiatives interact with each other and with children's own experiences and their family characteristics to influence childhood weight.”

Individuals Consulted on Statistical Aspects of the Design:

Warren Strauss  
Battelle Memorial Institute  
614-424-4275

Christopher Sroka  
Battelle Memorial Institute  
614-424-7876

Eric Leifer  
National Heart, Lung, and Blood Institute  
301-435-0419

Paul Sorlie  
National Heart, Lung, and Blood Institute  
301.435.0456

Colin Wu  
National Heart, Lung, and Blood Institute  
301-435-0440

Kevin Dodd  
National Cancer Institute  
301-496-7461

Chris Chapman  
National Center for Education Statistics  
202 502-7414

Contractors Responsible for Collecting Information for the Agency:

Battelle Memorial Institute  
505 King Avenue  
Columbus OH, 43201-2693  
Contact: Dr. Howard Fishbein  
(703) 248-1647