Evaluating Progress of Community-level Obesity Prevention Initiatives

January 5, 2013

Prepared by the
Center for Community Health
and Evaluation

Authors:
Allen Cheadle, Center for Community Health and Evaluation
Suzanne Rauzon, University of California Berkeley
Carol Cahill, Center for Community Health and Evaluation
Diana Charbonneau, Center for Community Health and Evaluation
Elena Kuo, Center for Community Health and Evaluation
Lisa Schafer, Center for Community Health and Evaluation

GroupHealth

www.cche.org

Abstract

Given the substantial resources being invested in community-level obesity prevention initiatives—multi-level, multi-sector, place-based efforts targeting the community food and physical environments—it is critical to assess their impact on behavior and health. However, there are many challenges to conducting credible evaluations of community-level initiatives, particularly evaluations that include an assessment of their impact on population-level behaviors and health outcomes. Longer-term outcomes such as improvements in food and physical activity behaviors are expensive and difficult to measure accurately at a population level and attribute to a multi-strategy initiative. Outcomes such as changes to the food and built environments are typically complex and multi-dimensional making it difficult to measure the extent of the environmental change and its likely impact on behavior. This paper reviews the range of methods that have been used or are currently being used to evaluate community-level obesity prevention initiatives and makes recommendations about best practices, innovative approaches, and the use of common metrics to assess the impact and improve the effectiveness obesity prevention initiatives.

Recommendations for implementing practical initiative evaluations include: (1) using logic model designs that link changes in the community environment to changes in population-level outcomes; (2) focusing evaluation resources on capturing behavioral and environmental changes and their potential impact; and (3) tracking longer-term population-level outcomes using primary or secondary data. More broadly, recommendations for evaluators of community-wide obesity prevention include: (1) increasing the evidence base on the impact of individual environmental strategies; (2) publishing the results of *all* evaluation studies, including negative ones; and (3) encouraging the widespread and systematic use of common validated measures that are practical to implement across settings versus in-depth, exhaustive measures that typically must be tailored to the objectives of an individual initiative.

Introduction

The obesity epidemic in the United States and elsewhere is driven by many complex and interrelated factors. Food and physical activity behaviors and weight gain are influenced by cultural, economic, social, genetic, and environmental influences that are hard to separate and even harder to change.[1-4] As a response to this multi-factorial challenge, an emphasis on comprehensive approaches to prevention have been developed, involving a portfolio of strategies at multiple levels (e.g., individual, family, community) across multiple sectors (e.g., school, worksite, neighborhood) following the socio-ecologic model or some modification of that framework.[5, 6]

Comprehensive community initiatives designed to address the multi-layered nature of many health risk behaviors date back to the early 1980s.[7-9] The approaches taken by these initiatives largely focused on implementing programs and carrying out social marketing campaigns. The limited success of these initiatives[10] led to a search for new approaches that were more comprehensive and sustainable, including an emphasis on policy and environmental change.[11, 12]

The environment—broadly defined to include the economic and social as well as physical environment—is a particularly promising area for obesity prevention. Place-based initiatives targeting the community environment have been singled out as promising intervention approaches in recent years.[13-17] The potential of environmental approaches was affirmed in consensus reviews by the Centers for Disease Control and Prevention (CDC) in 2009[18] and the Institute of Medicine (IOM**)** in 2006 and 2009.[19, 20] Examples of environmental approaches recommended for community-level initiatives include increasing the availability and affordability of healthier food and beverage choices in public service venues; increasing the geographic availability of supermarkets in underserved areas, restricting the availability of less healthy foods and beverages in public service venues, and enhancing the community infrastructure to support bicycling and walking.

Changing policies related to food and physical activity is another promising area receiving attention.[11] Examples of policy changes include sugar-sweetened beverage taxes and school policies limiting the availability of unhealthy food in vending machines and cafeterias.[21] It should be noted that passing policies does not by itself lead to the environmental changes that support improved behaviors—the policies must be effectively implemented. The advantage of policies is that they establish structure and guidelines that can lead to sustainable changes in the environment.

This emphasis on multi-level, multi-sector, place-based environmental approaches has been influential in shaping government and foundation initiatives aimed at preventing obesity, including the W.K. Kellogg Foundation's Food and Fitness Initiative,[22] the Robert Wood Johnson Foundation's Health Kids/Healthy Communities initiative,[23] the Kaiser Permanente Community Health Initiative,[24] the Department of Health and Human Services Communities Putting Prevention to Work Initiative funded under the American Relief and Reinvestment Act of 2009,[25] First Lady Michelle Obama's Let's Move Campaign,[26] and the White House Task Force on Obesity.[27] And passage of the Affordable Care Act includes prevention funding that is being used for place-based initiatives.[28]

Given the substantial resources being invested in "community-level initiatives" (multi-level, multi-sector set of activities focused on a defined community, including policy and environmental changes) to prevent and address obesity, it is critical to assess their effectiveness at changing health-related

behaviors. However, there are many challenges to conducting credible evaluations of community-level initiatives, particularly evaluations that include an assessment of their impact on population-level behaviors and health outcomes (e.g., obesity rates, diabetes prevalence). The key longer-term outcomes, especially food and physical activity behaviors, are difficult to measure accurately at a population level, and there are a number of challenges to creating designs and data collection strategies that can be sensitive enough to detect the relatively small expected changes and attribute them to the initiative. Intermediate outcomes such as changes to the food and built environments are typically complex and multi-dimensional, making it difficult to measure the extent of environmental change and the likely impact on behavior.

The purpose of this paper is to review the range of methods that have been used or are currently being used to evaluate community-level obesity prevention initiatives and make recommendations about best practices, innovative approaches, and the use of common metrics to assess the impact and improve the effectiveness of obesity prevention initiatives. After defining key terms, we will review evaluation goals and challenges; summarize the key issues related to design, data collection, analysis, and community engagement; and make recommendations for individual community-level evaluations and for the field as a whole. The goal of the paper is not to provide an exhaustive review of the community interventions or community evaluation measures, but rather to surface key issues related to evaluations of community obesity prevention initiatives, and make recommendations about how they might be addressed.

Definitions

Several terms are in wide use in a variety of different contexts. Clear definitions are required for a common understanding of their use related to initiatives that pursue impact on health behaviors and health outcomes. The following are brief working definitions we have developed based on the literature and common usage in the field. Two key terms are:

- **Community-level initiative.** For the purposes of this paper, a community-level obesity prevention initiative is defined as a multi-*level*, multi-*sector* set of activities focused on a *defined geographic community*, including policy and environmental changes, not solely programs serving individuals. *Levels* refer to levels in the socio-ecologic model (e.g., individual, family, organization, community). *Sectors* include organizational groupings (e.g., schools, workplace, health care, faith-based institutions), as well as a *community sector* that encompasses community-wide environmental changes (e.g. changes in social norms, community-level changes to the built environment). *Defined geographic community* means an area surrounded by known boundaries, such as streets or census tracts.
- **Community change strategies.** The primary goal of community-level obesity prevention initiatives is to bring about changes that will impact relevant behaviors, weight status, and health. Fawcett and colleagues[29] define a community change as a change in community programs, policies, or the environment. A community change strategy is a set of coordinated activities designed to lead to a sustainable community change.[30]

The community changes that are the target of community-level initiatives can be grouped into two broad categories—environmental change and programmatic change:

- **Environmental change.** Changes in the physical, economic, and social surroundings that have the potential to influence the behavior of a significant number of people (i.e., is not limited to an individual family/home environment).
- **Programmatic change.** Changes in the attitudes, knowledge, or skills of people exposed the program (e.g., classes, educational sessions, programmatic campaigns). Programs are interventions where participants (those exposed) and non-participants can be clearly identified. The programmatic focus in community-level initiatives is on implementation and maintenance of high-reach, evidence-based programs, where the programs are either integrated into existing organizations or sustained through long-term funding.

Making changes to the environment and programs can be enhanced by making changes to policies, systems and infrastructure, and by increasing community capacity. The following are definitions for these terms:

- **Policy change.** Changes in written procedures, laws, regulations, and, ideally, budget commitments designed to influence personal and organizational behavior in both public and private sectors.
- **Infrastructure change.** Sustainable changes in practices, procedures, and protocols within institutions and organizations designed to influence behaviors that do not involve formal written policies. Examples are establishing new communication channels between health and planning departments in a city government or instituting regular training in whole foods cooking for school food service staff.
- **Systems change.** Changes in the way organizations in different sectors interact with each other in an ongoing way; for example, strengthening the relationship between schools and city transportation departments around Safe Routes to Schools efforts. In contrast, changes within a single organization, e.g., one city agency or one school district, constitutes a policy or infrastructure change, not a systems change. The systems changes may or may not be supported by formal/written policy changes.
- **Capacity change.** Changes in organizational and community norms, resources, and relationships that increase the potential for implementing programs and sustainable environmental, policy, and systems changes. Examples include an increased awareness about the health impact of development within a city planning department, or the importance of whole foods cooking among school food service staff.

Note that policy changes are restricted to *written* policies and procedures; infrastructure changes are changes in procedures and practices within an organization (e.g., a school district) that are well embedded but not formally in writing. In both cases, it is important to count only those changes that might be expected to survive a change in organizational leadership.

Evaluation Goals and Challenges

Before reviewing current evaluation practices and potential innovations we provide a brief summary of evaluation goals and the challenges involved in evaluating community-level initiatives.

Evaluation goals

The primary evaluation goal for community-level initiatives is to assess their impact on food and physical activity behaviors, and longer-term impact on obesity rates and health outcomes related to obesity (e.g., diabetes, heart disease). Substantial investments are being made in these initiatives and there is considerable interest in understanding long-term impacts and the value of the investments. However, evaluation can also play a formative role in program development. Information gathered as part of the evaluation is often useful in program improvement, particularly more proximal process information related to the implementation of community change strategies. For example, lessons may be learned and communicated to community coordinators regarding working more effectively with local decision makers to implement built environment strategies or with schools about formulating food and physical activity policies. Evaluators may also provide guidance around which strategies are most effective and likely to impact behavioral and health outcomes during the planning phase.

Evaluation challenges

The challenges associated with conducting useful evaluations of community-level initiatives can be grouped in two broad areas: (1) measuring changes in the community environment and determining their potential impact on behavior change; and (2) measuring changes in longer-term outcomes (behavior, obesity, health outcomes) and attributing the changes to community-level initiatives.

Measuring changes in the community environment and rating potential impact

Characterizing/quantifying environmental change. Environmental changes are typically complex and multi-dimensional, even for relatively simple and straightforward environmental interventions. This complexity makes it difficult to create indices that can be used to make comparisons about the degree of environmental change, either over time within a single community or across multiple communities. For example, a built environment strategy in one community may focus on extending sidewalks to more parts of a neighborhood, while another community may add bike lanes and traffic calming interventions to increase safety for cyclists and walkers. Both strategies are attempting to increase walking, but how can they be placed on a common yardstick that measures the degree of environmental change? Or one school cafeteria strategy may increase the number of healthier entrees and reduce unhealthy ones, while another school only adds a salad bar. The complexity and varying environmental change goals also makes it difficult to develop a core set of common measures for environmental change. And the presence, or lack thereof, of complementary supporting strategies (educational, programmatic, promotional), can greatly influence the degree of change.

Documenting implementation. A second measurement challenge for both environmental and programmatic changes is determining the extent to which the changes have been implemented. This is particularly true for changes that occur in large institutions (schools, large worksites) and are decentralized across the organization. For example, one school strategy is to reduce the extent to which teachers use candy as a reward for classroom behavior. Self-report data at the district or school administrative level is often flawed since school administrators often do not know details about what

each teacher is doing their classroom. And surveying teachers can be resource intensive and burdensome, and relies on self-reported data that may be influenced by a social desirability bias. The only certain way to determine whether a policy is being followed is direct observation over repeated visits, which is not usually feasible.

Rating potential impact. Another challenge is estimating the impact on behavior from particular environmental change strategies. For example, can we expect more increases in overall minutes of physical activity to result from extending the sidewalk network, improving traffic safety, or implementing more sports programs at a neighborhood community center? Meaningful answers require knowing the magnitude of the effects of different interventions and the potential to reach the intended audience. Ideally, these estimates would come from research evidence published in peer-reviewed reports; however, despite the strong support for environmental approaches in the CDC and IOM reports,[18-20] there is a lack of empirical evidence about the degree to which increasing the availability of more healthful food promotes healthier eating or whether changing the built environment leads to increases in physical activity. The CDC recommendations were based solely on expert panel ratings of strategies that have been mentioned prominently in the literature. The IOM panel on local government action assembled a wide range of literature and reports in support of its recommendations, but did not attempt a systematic meta-analysis of the available intervention studies. The CDC Community Guide,[31] which does rely on available evidence in making recommendations, has no recommended strategies involving environmental change for either nutrition or obesity. One reason for the lack of evidence is that the field is relatively new, though results from environmental change intervention studies are appearing with increasing frequency. What Works for Health, a resource associated with the County Health Rankings model of assessing community needs, lists a number of environmental change strategies with evidence of effectiveness.[32] However, most studies of environmental impact continue to be cross-sectional studies, for example, related to healthy food availability in neighborhoods[33, 34] or changes to the built environment for increasing physical activity.[35-37]

<u>Measuring and attributing changes in longer-term population-level outcomes</u>

Obtaining valid, reliable outcome data. All of the key longer-term outcomes for community-level initiatives are difficult to measure at the population-level with the evaluation resources typically available. Food and physical activity behaviors are complex and multidimensional, and the "gold standard" research methods for capturing key measures (e.g., food and caloric intake patterns and minutes of physical activity) are labor intensive and time consuming to collect and analyze (e.g., food diaries and accelerometer studies). Self-reported measures using brief phone and/or paper surveys (e.g., the national Behavioral Risk Factor Surveillance System (BRFSS); see a description in the following section on data collection) have been shown to have reasonable validity among large samples of adults (e.g., servings of fruits and vegetables, minutes of leisure time physical activity), but are often not sensitive enough measures for localized community interventions at the county or neighborhood level. Moreover, self-reported measures are less valid and reliable among children, which are often the primary target of obesity prevention efforts.

Body mass index (BMI), derived from measures of height and weight, is a reliable measurement technique, but expensive to collect outside of health care clinical settings (e.g., in school clinics). Also, rigorous protocols and well-trained staff are needed to obtain accurate and consistent height and weight measurements.[38] The desire to obtain population-level estimates of BMI to track long-term

impact often leads to the collection of less-accurate self-reported height and weight information from subjects across the entire population.

Detecting modest population-level changes. The promise of community-level interventions is that even small effects on each person can result in a large overall impact in terms of population-level weight status or health care cost savings. Simulation models illustrate this impact:[39] Even modest changes in the trajectory of obesity can have a sizable long-term impact on population-level rates of diabetes and heart disease**.** Furthermore, small changes are likely to be expected given the modest scope of community-level interventions relative to the array of factors that shape physical activity and dietary behaviors.[10, 40] Unfortunately, the small population-level changes expected (and hoped for) are difficult to detect given the measurement and sampling error associated with population-level surveys.[10, 40] Achieving the sample sizes required for adequate power to detect changes is difficult given the cost of population-level surveys.[41] And it is difficult to obtain response rates that are representative of the entire population of a community without a substantial investment in multiple contacts to obtain completed surveys.[42, 43]

Attributing observed population-level changes to the initiative. Even if population-level improvements in key indicators are observed, how likely is it that they are due to the community-level initiative being evaluated? Because of the high degree of attention nationwide to the obesity epidemic, many changes are under way at all levels—local, state, and national. A place-based initiative often has little control or knowledge of these changes outside the scope of its action plans. The private sector is responding to some degree (e.g., a few in the food industry voluntarily posting calorie information on restaurant menus), which complements the numerous public sector initiatives. Given the large number of overlapping efforts, true comparison communities are very difficult to locate, and may not exist, making conventional quasi-experimental designs less realistic.

Community Initiative Evaluation Methods

This section reviews the design, data collection, and analysis methods that have been used or are currently being used in evaluating community-level obesity prevention initiatives. A review of the published and "grey" literature (e.g., unpublished evaluation studies and online initiative descriptions), was conducted to identify examples of initiatives that have been or are currently being evaluated. A limitation to this review is that despite the rapidly growing interest in community-level approaches to obesity prevention, there are relatively few published studies of community-level initiatives targeting children and adults with multi-level, multi-sector interventions. Most of the studies we were able to locate focused on children rather than adults or families. And the most rigorous studies tended to be intensive interventions with an emphasis on school-based programs (and thus not truly "multi-sectoral"), although some of these programs had supporting community-level components.

A search was conducted in December 2012 for years 2000-2012 using bibliographic databases, notably the National Library of Medicine's PubMed, and web sites that aggregate reports on obesity prevention interventions, such as the Agency for Healthcare Research and Quality (AHRQ) *Innovations Exchange* and the Robert Wood Johnson Foundation's *Active Living Research* program. A total of 36 community-level initiatives that included sufficient detail concerning their intervention and evaluation methods were identified (see Appendix A for a list of the initiatives along with brief descriptive information). These included 16 initiatives that were completed and included population-level outcome results (3

negative studies, 13 positive). Another 20 initiatives are either in process or not measuring population-level behavior change. Some of the largest and potentially most useful evaluations are in progress. In particular, many independent evaluations of the CDC's Communities Putting Prevention to Work initiatives that are currently being conducted, and a large-scale, retrospective NIH-funded Healthy Communities Study that is using chart reviews to track changes in obesity rates in 268 communities across the United States.[44] Perhaps the best known of the completed studies is Shape up Somerville, a comprehensive community-level intervention, involving children, parents, teachers, school food service providers, city departments, policy makers, health care providers, before- and after-school programs, restaurants, and the media.[45] The Shape up Somerville interventions resulted in a modest, but significant decline in BMI z-scores in children in grades 1-3.

Evaluation Designs

Broadly speaking, the strongest evaluation design is one that will give the most credible answer to the question: "Does *a given* community-level initiative have an impact on population-level food and physical activity behaviors and longer-term outcomes such as obesity rates?" The words "*a given*" are italicized to emphasize that it is probably unrealistic to expect that one can generalize from a single evaluation or research study. In an ideal world, several large-scale, multi-community experimental studies that are generalizable to a larger group of communities would be conducted. At one point, beginning with the earliest community-level health promotion intervention studies in the 1980s,[7-9] there was an attempt to conduct such experimental studies. For the studies to be truly generalizable, there must be random selection of a representative set of communities, both for initial inclusion in the study and for assigning to intervention and control conditions. In particular, it requires selecting communities with a range of readiness/capacity to successfully carry out a community-level initiative, including both low and high capacity communities. However, given the substantial resources required to carry out a community-level initiative, funders have been understandably reluctant to choose random/representative communities to participate, and instead fund those that appear most likely to be successful. It is still possible to conduct randomization or quasi-randomization once a group of communities is selected, and a number of evaluations[7, 9] have done so. But evaluation design challenges, especially study power given the small number of communities that can be funded and the scope required to collect useful data, have resulted in few evaluations with experimental or quasi-experimental designs that have shown conclusive results or an intervention effect.[10] Additionally, it is often difficult to find communities that are willing to serve as a "control" group and forego services and programs even if they are not yet of proven benefit.

As a result of the funding constraints and design challenges, the field has largely moved away from attempting to design large-scale experimental studies to determining whether a particular community-level intervention worked. There are still group randomized trials at levels less aggregated than community (e.g., school, worksite) and at least one retrospective study currently underway— the Healthy Communities Study[44]—that includes large numbers of communities across the United States, selected with rigorous protocols and sampling design. However, most evaluation designs are similar to Shape up Somerville, most often cited as a successful community-level initiative, which was conducted in a single community with two comparison communities.[45] Almost all of the reviewed evaluations were in one community or a small number of communities. Focusing the design on determining the impact of a particular initiative is a reasonable strategy given all of the challenges to large-scale, multi-site trials, but it does change the long-term approach to building the evidence base about which community-level

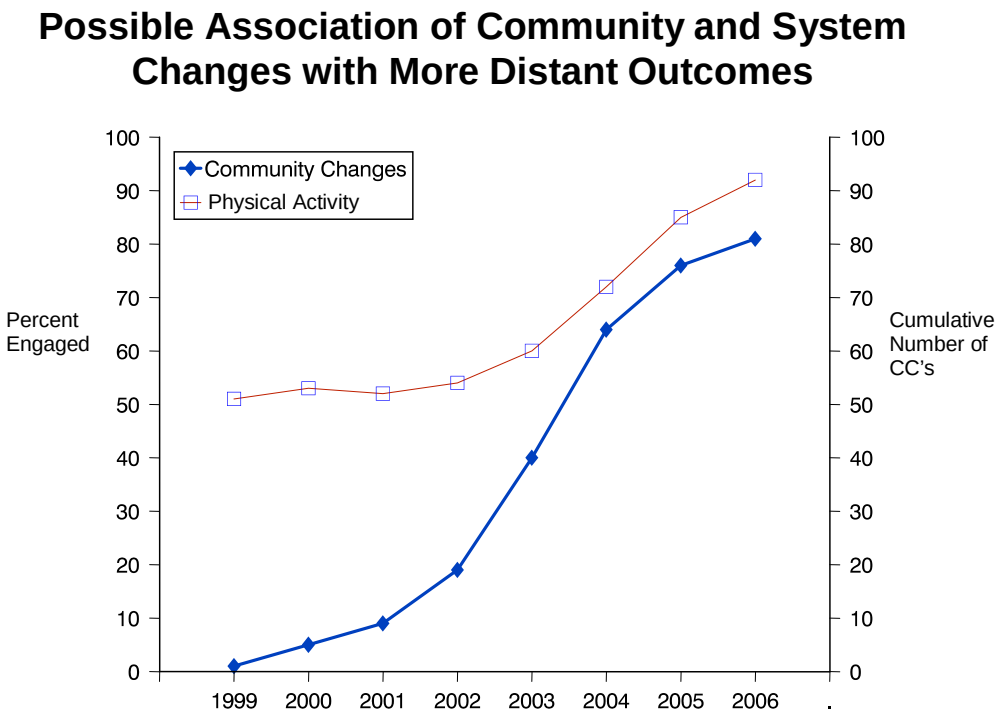
intervention approaches work and why. In particular, an accumulation of results from many evaluations of individual strategies rather than a handful of large-scale experimental or quasi-experimental studies will be required (see the section about building the field recommendations).

Assuming that the desired focus is determining the impact of a given community-level initiative, there are two evaluation design choices—"conventional" designs and "logic model" designs. Conventional designs are typically quasi-experimental where a set of outcome measures are specified in advance and measured pre/post in both intervention and comparison communities, although they can include pre/post measurement without comparisons. Of the 16 completed evaluations with population-level results all but one were conventional designs; 9 were quasi-experimental, and 6 were pre/post without comparisons (See Table A-1, Appendix).

Logic model designs[46] start with a program "theory of change," i.e., the mechanism by which the comprehensive community initiative is intended to achieve its long-term outcomes, and then create indicators for each step in the logic model. In the case of community-level initiatives, the key steps in the logic model are intermediate outcomes (e.g., environmental and policy changes implemented in communities) and longer-term population-level outcomes (e.g., physical activity behaviors, weight and health status). If the temporal pattern of change is consistent with that specified in the logic model, the intervention is more likely to have been the cause of the population-level changes; for example, if significant built environment changes are made to promote walking and the time trend in minutes of daily walking shift upward in the years that follow.

Examples of two approaches that systematically apply logic model designs to community-level initiatives data are methods developed by Fawcett and colleagues[47] and the "population dose" concept developed as part of the evaluation of the Kaiser Permanente Community Health Initiative (KP CHI).[30] Fawcett creates an index of the degree of community change taking place by counting up the number of community changes (i.e., changes in programs, policies, practices, and the environment) over time in a given community. An interrupted time series approach is then used that overlays a plot of the cumulative community changes and with a plot of the trend in a population-level outcome (e.g., behavior change). See Figure 1 for an illustrative figure (drawn from a handout developed by the Kansas Work Group).

Figure 1. Example of Kansas Work Group Attribution Approach



If shifts in the population-level outcome trend line coincide temporally to shifts in the trend of community changes then it is plausible to attribute the population-level changes to the community-level initiative. This method has been used successfully by Fawcett and colleagues in a number of initiatives.[48]

The "population dose" approach uses elements of the RE-AIM method of combining reach and effectiveness[49] to estimate the likely impact of a community change strategy on population-level behavior. Population dose is defined operationally as the product of penetration (reach divided by the size of the target population) and effect size (relative change in behavior for each person exposed, e.g., 10% increase in minutes walked per day among residents living near a newly installed walking trail). For example, if 20% of the community target population lives near a new walking trail and the average effect size is 10% for each person exposed (living near the trail), the population dose is 20% x 10% = 2%. Essentially, population dose is the effect size of the intervention, if the effect was spread across all of the residents of the target community. Since quantitative effect sizes for policy and environmental change interventions are generally unavailable in the literature, this method uses a three-level rating system (high/medium/low) to assess the strength of most intervention strategies; methods are described elsewhere.[30]

The dose ratings are then combined with population-level data to examine whether higher dose community change strategies or clusters of strategies are associated with measured population-level changes in the relevant outcomes. For example, if a number of built environment changes are rated as high dose for promoting walking, then a survey of community residents should show measurable increases in minutes walked. In the first test of the dose approach, 9 instances were identified where there were high-dose strategies or clusters of strategies targeting particular outcomes in 3 KP CHI communities in Northern California. All 9 high-dose strategy clusters were targeting youth in schools;

none of the neighborhood-based strategies targeting adults or families were rated high dose. Then, pre/post changes in the relevant youth survey measures were compared for these 9 high-dose strategy clusters. Positive and significant findings for 4 out of 9 comparisons were found.[50]

The advantage of logic model designs is that they are more "specific," i.e., better able to rule out false positives where a favorable population-level change occurred that was not the result of the initiative. Since there are often multiple outcomes being measured (e.g., food behaviors, physical activity behaviors, obesity rates), typically at least some of the outcomes will show positive results. With a conventional design, beyond applying adjustments for multiple comparisons, there is little that can be done to sort out the underlying cause of the positive changes. However, in a logic model design, if a behavioral outcome improves but there are no corresponding community changes or high-dose intervention, it is much harder to conclude the initiative was responsible for the positive outcome change. The disadvantage of the logic model approach is that it requires an accurate assessment of changes in environmental outcomes or dose, which can be challenging as noted previously.

Data collection

Best practices in data collection are a balance between rigorous, detailed assessments and less rigorous, less precise measures that are cost effective and can be accomplished quickly. Of primary interest is data capturing environmental change outcomes and their potential impact; and longer-term population level outcomes.

Process evaluation/documenting implementation. There are a number of ways to track the process of implementing strategies in community-level initiatives, ranging from detailed, real-time reporting to retrospective progress reports at the end of an initiative. Regular (monthly or quarterly) reporting is most effective for generating accurate data while at the same time building relationships with community intervention staff that aids communications about potential intervention improvements. This ongoing reporting can be web-based (e.g., the Community Tool Box system developed by Fawcett and colleagues), or over the phone (such as the monthly calls made as part of the KP CHI evaluation). The reporting should focus primarily on implementation of community change strategies and success factors/barriers related to implementation.

Intermediate outcomes. Intermediate outcomes such as environmental and policy change are at the core of the community-level initiative approach, so there has been a substantial effort directed at developing and validating tools, making them available through online inventories (see Table 1 for a partial list), and reviewing and making recommendations about measurement approaches.

Table 1. Tools for Evaluating Community Obesity Prevention Initiatives

Source	Description	Website
COLLECTIONS		
Active Living Research	Tools to collect data on streets, schools, parks, or other community settings to see how well they support physical activity.	http://www.activelivingresearch.org/toolsandresources/toolsandmeasures
National Collaborative on Childhood Obesity Research (NCCOR) Measures Registry	Searchable database of diet and physical activity measures relevant to childhood obesity research. Measures included describe, monitor, and evaluate interventions, particularly policy and environmental interventions, and factors and outcomes at all levels of the socio-ecological model.	http://tools.nccor.org/measures/
National Cancer Institute Risk Factor Monitoring & Methods	Tools for researchers, including dietary surveys and environmental assessments	http://riskfactor.cancer.gov/
SELECTED ENVIRONMENT MEASUREMENT TOOLS		
Environmental Assessment of Public Recreation Spaces (EAPRS)	Comprehensive direct observation assessment of the physical environments of parks and playgrounds, with an emphasis on evaluating physical elements and qualities with respect to their functionality or potential functionality (e.g., how a park or playground element is used or could be used by adults and children).	http://www.seattlechildrens.org/research/child-health-behavior-and-development/saelens-lab/measures-and-protocols/
Irvine Minnesota Inventory	Measures a wide range of built environment features that may affect physical activity, especially walking. It includes 160 items, which covering 4 domains: accessibility, pleasurability, perceived safety from traffic, and perceived safety from crime.	https://webfiles.uci.edu/kday/public/index.html
Nutrition Environment Measures Survey (NEMS)	Measures focus on surveying community and consumer nutrition environments, which include the type and location of food outlets (stores and restaurants), availability of healthful choices and information, pricing, promotion, and placement of healthier food products.	http://www.med.upenn.edu/nems/
Communities of Excellence in Nutrition, Physical Activity & Obesity Prevention (CX3)	Field surveys of neighborhood food access.	http://www.cdph.ca.gov/programs/cpns/Pages/CX3_T2_FieldSurveys.aspx

SELECTED POLICY MEASUREMENT TOOLS		
Bridging the Gap Research Informing Policy and Practices for Healthy Youth	Includes surveys of school district policies and practices related to childhood obesity and tools for coding school district wellness policies.	http://www.bridgingthegapresearch.org/research/district_wellness_policies/
University of California, Berkeley Center for Weight and Health Evaluation/Tools	Surveys include Nutrition Learning Environments, Actions, & Policies (Nutrition LEAP); Nutrition Services Questionnaire; and Survey of Child Care Providers.	http://cwh.berkeley.edu/center/evaluation_tools
WellSAT: Wellness School Assessment Tool	Online tool for quantitative assessment of school wellness policies; from the Yale Rudd Center for Food Policy & Obesity.	http://www.wellsat.org/
School Health Index	Center for Disease Control's online self-assessment and planning tool that schools can use to improve their health and safety policies and programs.	http://www.cdc.gov/healthyyouth/shi/index.htm
CoalitionsWork Tools & Resources	Resources include assessments of community and state plans for obesity prevention.	http://coalitionswork.com/resources/tools/
TRAINING		
BEAT (Built Environment Assessment Training) Institute online training	Free courses on assessing the built environment for physical activity, including an in-depth look at specific tools, and assessing the nutrition environment with the Nutrition Environment Measures Survey (NEMS).	http://www.med.upenn.edu/beat/onlinetraining.shtml

A special issue of American Journal of Preventive Medicine in 2009 was devoted to measuring the food and physical activity environment, which included a summary highlighting key considerations in tool/instrument development.[51] Saelens and Glanz provide a good summary of the central issue in environmental measurement, namely the trade-off between having a single instrument measuring an environment (e.g., parks, grocery stores) and having multiple instruments to capture the diversity of intervention approaches: "Multiple instruments of the same or similar-enough constructs make it difficult to compare across studies or time to help derive generalizable estimates of association or effect. This situation likely results from a lack of awareness of existing measures, from the concurrent development of similar measures without outlets and incentives for ongoing dialogue for work in progress or dissemination, and/or the belief that an existing instrument fails to capture the construct adequately (e.g., instrument not perceived as applicable in a certain geographic area or with a given population). Physical activity and eating behaviors have considerable diversity, so it is not surprising that instruments that attempt to measure their respective environments are diverse" (p.s166-7).

One nuance related to the diversity of existing measurement tools and efforts to standardize is the distinction between *tools* and *measures*. *Tools* can be defined as instruments, often long, that provide a

relatively comprehensive assessment of the environment within a setting or institution (e.g., park, grocery store, restaurant). *Measures* are a subset of items from a longer tool and focus on measuring a particular construct (e.g., the characteristics of the walking trails within a park, the number and quality of fruits and vegetables in a grocery store). The literature we have found tends to report on inventories of tools, rather than measures. Given that initiatives often focus on different aspects of an environment (e.g., some grocery store strategies focus on increasing the availability of produce, others on reducing unhealthy snacks), it is very difficult to create a single, standardized tool that can capture all potentially relevant aspects of an environment and still be of reasonable length for administration. However, there is much greater likelihood that a single measure could be identified that would provide the most economical and valid way of representing a construct. Customized tools could then be developed for different initiatives drawing on a standard library of measures.

Another observation related to environmental measurement is that there may be instances when doing pre-post environmental assessments using standardized tools is not the preferred approach. Environmental initiatives often change course over time, switching venues (e.g., going to a different grocery store) or the targets within an institution (e.g., changing from an emphasis on healthy snacks to increasing the supply of fresh produce). Tools that were appropriate at baseline may not capture the key changes at follow-up. Furthermore, standard tools are often long and time consuming to implement, require extensive training in data collection, and they may obtain information on environmental dimensions that are not the ultimate target of an initiative. One alternative, particularly with evaluations with fewer resources, is to do a retrospective assessment using key informants (e.g., community coordinators, grocery store managers) to report on the changes that were made, perhaps supplemented with photos. A disadvantage is that the resulting data are a series of qualitative descriptions across institutions and communities of environmental changes that must be categorized and rated as to their likely impact on behavior change.

Finally, there are useful sources of secondary data on the food and physical activity environment. Where initiatives are targeting environmental dimensions related to these data, for example, increasing access to farmers markets and healthy food stores or increasing neighborhood safety to promote outdoor physical activity, the data can provide a low-cost alternative to primary data collection and also provide comparison communities and a longer time frame of available data. There are also national policy databases that can be useful in tracking policy implementation. Table 2 gives a sample listing of these secondary data sources.

Table 2. Selected national datasets for evaluating policy and environmental changes from community health initiatives

Focus	Resource	Source	Description	Geography	Website
ENVIRONMENTS					
Nutrition	Food Environment Atlas	USDA	Compilation of data from various federal agencies including access to grocery stores, restaurants, recreation facilities.	County	http://www.ers.usda.gov/data-products/food-environment-atlas.aspx
Nutrition	Food desert locator	USDA	Low-income census tracts where a substantial number or share of residents has low access to a supermarket or large grocery store.	Census tract	http://www.ers.usda.gov/data-products/food-desert-locator.aspx
Nutrition	Farmers Market Directory	USDA	A searchable database & data download.	Street address	http://search.ams.usda.gov/farmersmarkets/
Nutrition	Supplemental Nutrition Assistance Program (SNAP) retail food locator	USDA	An interactive map & data download.	Street address & latitude/ longitude	http://www.snapretailerlocator.com/
Nutrition	Supplemental Nutrition Assistance Program (SNAP) participation	USDA	Percentage of population that receives SNAP benefits.	County	http://www.ers.usda.gov/data-products/supplemental-nutrition-assistance-program-(snap)-data-system.aspx
Physical activity	Uniform Crime Reporting System	FBI	Compilation of crime statistics (violent crime/property crime).	Law enforcement jurisdiction	http://www.ucrdatatool.gov/ranking.cfm
POLICY					
Nutrition & physical activity	Classification of Laws Associated with School Students (C.L.A.S.S.)	NCI	Searchable database. C.L.A.S.S. uses a scoring system to classify state laws as they compare to national standards and recommendations for PE and nutrition; available at elementary, middle and high school levels.	State	http://class.cancer.gov/

Notes:

CDC: Centers for Disease Control & Prevention

FBI: Federal Bureau of Investigation

NCI: National Cancer Institute

USDA: U.S. Department of Agriculture

Population-level Measures. The issues related to population-level measurement center around a series of trade-offs driven by the size of the target community and available evaluation resources. The trade-offs include: (1) using primary versus secondary data sources, (2) the length of time over which population-level data will be tracked, and (3) the comprehensiveness and validity of the measures used. Before examining those trade-offs, here is a brief summary of primary data collection methods and secondary data sources.

Table 3 summarizes the principal population-level, primary data collection options for youth and adults. Given the proliferation of cell phones, mail surveys (often supplemented by phone and web follow-up) are replacing random-digit dialing as the method of choice for surveys of adults. However, response rates continue to decline for all surveys.[42, 43] Opinion and political polls are increasingly using cellular phone numbers, but with the advent of Local Number Portability (LNP), a system that enables end users to keep their telephone numbers when switching from one communications service provider to another, these numbers are not reliably linked to a place of residence. Clinical data are becoming a more viable option as Health Data Repositories and Health Information Exchanges are being created that pool electronic health records of multiple health care providers within a given geographic area.

Table 3. Population-level measurement method options for youth and adults

Data collection methods	Comments
ADULTS	
Phone surveys • Random digit dialing • List-based	• Cell phones make it increasingly difficult to obtain representative samples for both list and random-digit dialing
Mail surveys • List-based with phone follow-up using reverse directories	• Becoming the state of the art for large-scale surveillance surveys • May be combined with a web/online option
In-person/door-to-door surveys	• Useful when interventions target smaller geographic areas or institutions such as low-income housing or apartment complexes
Interactive voice response (robo calls) • List based or RDD	• Low response rates (5%-10% in the Kaiser Permanente Community Health Initiative) • Cost per completed survey about half of phone and mail surveys
Online surveys	• Difficult to secure a representative sample of email addresses
Clinical data • Body mass index • Behavioral measures	• Can be used when clinic penetration in a community is high
YOUTH	
School-based • Paper and pencil surveys • Electronic (tablet) surveys	• Most efficient method • Can be difficult to secure school participation • Using electronic devices for survey administration (e.g., tablet computers) can make data management more efficient • Not appropriate for children younger than 4th/5th grade

Data collection methods	Comments
Fitnessgram (http://www.fitnessgram.net/home/) • Body Composition (Height, weight, BMI or percent body fat) • Aerobic capacity • Muscular strength, flexibility, endurance	• Provides multiple measures of fitness, weight or body composition status
Proxy surveys • Parents responding for children	• Used in California Health Interview Survey

Table 4 provides a summary of national secondary data sources for local areas, including those most relevant for obesity prevention, but a number of others that can be useful for adjusting for socio-demographic factors.

Table 4. Selected datasets with population-level data for evaluating community health initiatives

Focus	Resource	Source	Description	Geography	Website
COLLECTIONS					
Health determinants	County Health Rankings	U of Wisconsin	Data compiled from various sources that correspond to a model of population health that emphasizes the many factors that, if improved, can help make communities healthier places to live, learn, work and play.	County	http://www.countyhealthrankings.org/our-approach
Health determinants	CHNA.org	Institute for People, Place & Possibility	Maps displaying indicators for local areas, compiled from various sources.	County & census tract	http://www.chna.org/
HEALTH BEHAVIORS & OUTCOMES					
Diet & physical activity; BMI & diabetes	Behavioral Risk Factor Surveillance System (BRFSS)	CDC	Telephone survey conducted year around. For most states and counties, BRFSS is the only source of population-based health behavior data related to chronic disease. Administered by states.	County; large urban health departments produce reports for smaller areas & populations	http://www.cdc.gov/brfss/
Diet & physical activity; BMI	Youth Risk Behavior Surveillance Survey (YRBSS)	CDC	Bi-annual survey of 9th-12th grade students. Administered by states and school districts; some states, (e.g., WA), have their own surveys.	State & some large school districts	http://www.cdc.gov/healthyyouth/yrbs/index.htm

Focus	Resource	Source	Description	Geography	Website
Diet & sedentary activity; BMI	Pediatric Nutrition Surveillance System (PedNSS)	CDC	From health departments that choose to participate submit data to CDC on a monthly basis. The number of PedNSS contributors differs slightly from year to year because all state health departments do not submit data every year.	National, state	http://www.cdc.gov/PEDNSS/
Diet & physical activity; BMI, lipid profile	National Health and Nutrition Examination Survey (NHANES)	CDC	Complex, stratified, multistage, probability sample of 10,000 individuals representative of the civilian, non-institutionalized US population. Oversampling of specific age groups, racial/ethnic groups, and low-socioeconomic status populations. Every 2-years.	National; state, county, census block (restricted use)	http://www.cdc.gov/nchs/nhanes.htm
Physical activity; health determinants	Commuting characteristics (means of transportation to work, including walking & bicycling); socio-demographic characteristics	Census Bureau	Five-year estimates from continuous American Community Survey; 2007-2011	Census tract	http://factfinder2.census.gov/faces/nav/jsf/pages/index.xhtml
Physical activity; obesity & diabetes	CDC - diabetes/ physical activity/ obesity	CDC	County-level estimates of diagnosed diabetes, obesity, and physical inactivity (BRFSS data).	County	http://apps.nccd.cdc.gov/DDT_STRS2/NationalDiabetesPrevalenceEstimates.aspx
Nutrition & physical activity	Market research reports	The Nielsen Company and various other sources	Numerous private firms produce fee-based market research reports. Example: Nielsen Homescan Data, which produces detailed food purchasing information from a panel of U.S. households.	State, county, metropolitan market area	http://en-us.nielsen.com/

Focus	Resource	Source	Description	Geography	Website
Physical fitness, BMI	Fitnessgram	Various States, Depts of Education	Collects body composition (height, weight, BMI or percent body fat, aerobic capacity, muscular strength, flexibility, endurance) from school-age children. Currently in approximately 67,000 schools in all 50 states. About 22 million students were tested in 2011. Largest states and district administering Fitnessgram testing are: Texas, Georgia, Delaware, North Carolina, Kansas, California, New Jersey. Another 25 states are over 50% covered.	State, District, School	California example: http://www.cde.ca.gov/ta/tg/pf/

Notes:
CDC: Centers for Disease Control & Prevention
FBI: Federal Bureau of Investigation
NCI: National Cancer Institute
USDA: U.S. Department of Agriculture

Data across the U.S. often are available only at the county level, which may not be that useful for evaluating initiatives targeted at small neighborhoods. Often, state or local health departments will be able to provide more granular data by request. Other jurisdictions, such as police departments, often report rates as well as numbers of events. Some states administer their own risk factor surveys, e.g., the California Health Interview Survey and the Healthy Youth Survey in Washington State.

Regarding the trade-offs noted above, the first concerns the use of primary versus secondary data. High-quality surveillance data, such as the BRFSS at the national level (Table 4), is clearly the preferred option, if feasible—the data are often freely available for download and provide an extended time series that includes comparison geographies. The principal drawback is sample size—most of these surveys are designed to provide results at the state or national level, and those that provide county-level data often only have adequate sample sizes for relatively large counties. Community-level initiatives usually target geographic areas much smaller than county, e.g., city or neighborhood, and therefore the impact of these initiatives might not be large enough to be detected by these types of surveys. An additional drawback is the often large time lag between data collection and public availability of the data. For example, as of January 1, 2013, the CDC had posted BRFSS data through 2011.

The second trade-off concerns the length of follow-up. Given the entrenched nature of the factors contributing to the obesity epidemic, it is reasonable to expect that community environmental and programmatic changes must be sustained over a long period of time for significant population-level impact to occur. However, most primary population-level data collection is constrained by the funding

period of the initiative, with the data endpoint coinciding with or coming soon after the intervention funding endpoint. To have a reasonable chance of detecting longer-term changes, some of the data collection resources must be shifted to one or more years beyond the initiative period, which means reducing sample sizes on each data collection occasion or finding more inexpensive, less comprehensive methods that permit a larger overall number of surveys over longer periods of time. One advantage of logic model designs described previously is that they can help focus long-term follow-up data collection only in those communities where the dose or extent of community changes suggests there is likely to be an observed impact.

The third trade-off concerns the degree of comprehensiveness and validity of the data collection methods. For example, the gold standard for food consumption behavior measurement is the 3-day food diary. Gold standard methods can be time-intensive and expensive to implement and often are not feasible given evaluation budget constraints. Alternatives are brief, self-report, frequency questions such as the fruit and vegetable serving questions on the BRFSS (e.g., how often do you eat green salad?). The issue is whether to invest data collection resources in collecting high-quality data in a smaller number of communities (or a smaller sample size per community), or to have less valid measurements in a larger sample. Options for compromise might be to use multiple methods in a handful of communities as a check on the validity of the less intensive measures or to conduct validity/reliability testing on self-report instruments.

Analysis

The analytic approaches for conventional evaluation designs are relatively straightforward. As stated above, the ideal study design would be a large group-randomized trial, and with it would come a gold standard analysis in the form of logistic regression or ANCOVA.[52] However, for the more typical single community pre/post evaluation design, t-tests and chi-square tests are often sufficient for measuring change. For multi-community initiatives and quasi-experimental studies with non-random comparison groups, adjustment must be made for clustering by community. In either case, if evaluators were able to collect such data, regressions can be conducted to control for demographic differences or other confounding influences over time or across communities. The large number of potential outcome variables requires some adjustment— either through Bonferroni corrections,[53] Sidak, Tukey, Dunnett's or other methods of adjusting for multiple measures.[54] As noted above, the problem of multiple comparisons is a significant problem for the interpretation of community level initiative results.

The analytic methods for logic model designs are less developed and less straightforward. Ultimately, the preferred approach is a regression model with community as the unit of analysis, behavioral measures and BMI as the dependent variables, and dose or some measure of likely impact as the independent variables. This is the approach planned for the Healthy Communities Study, and the large sample size (n=268 communities) aims to give adequate power to detect significant relationships between intervention strength and outcomes. For evaluations of a small number of community initiatives, the analytic approach can follow what was referenced above in the KP CHI, namely (1) identifying the dose of the interventions related to each of the major outcome variables (e.g., minutes of physical activity, servings of fruits and vegetables, consumption of unhealthy snacks); and (2) verifying whether the pattern of outcome changes follows the pattern of intervention dose, applying tests of significance such as sign tests if there are a large enough number of comparisons.

Community engagement and participation in evaluation

Community engagement in health promotion research and evaluation covers a wide continuum of practice that goes by many names: community engaged research, community based participatory research, action research, empowerment evaluation, participatory evaluation, transformative evaluation. Other disciplines, including sociology, political science, cultural anthropology, organizational development, psychology, and social work, have all contributed to the field.[55] CDC defines community engagement as "the process of working collaboratively with and through groups of people affiliated by geographic proximity, special interest, or similar situations to address issues affecting the well-being of those people."[56] This section reviews the advantages and disadvantages of community engagement in evaluation and commonly used approaches to engagement. Recommendations about the appropriate level of engagement are provided in the overall recommendations section.

Table 5 summarizes the advantages and disadvantages of engaging communities in evaluation. Advantages include better interventions, more robust methods and results, and greater trust in the evaluation results by communities, which may make them more likely to adopt recommendations for program improvement. Disadvantages include time burden on community members and a lack of skill in community engagement on the part of many evaluators.

Table 5. Advantages and disadvantages of community engagement in research and evaluation

Advantage/Disadvantage	Rationale
ADVANTAGES	
Promotes ownership in the intervention development process	Interventions can achieve better outcomes for health by promoting community "ownership" in the process.[57]
Addresses health disparities	If health inequalities are to be adequately addressed, approaches to health improvement must take into account the concerns of communities and be able to benefit diverse populations.
Improves the quality of the research/evaluation design	Collaborative evaluation approaches can provide a more comprehensive understanding of factors that need to be accounted for in the design.[58]
Increases response rates, improves data quality	Partnering with community organizations and engaging residents can increase the visibility of the evaluation, increasing response rates and facilitating access to key informants.
Contributes to the relevance and usefulness of evaluation data	Stakeholder involvement in evaluation is potentially powerful in developing program practitioners' sense of ownership and understanding of programs and can lead to conceptual and instrumental uses of evaluation data.
Contributes to a more nuanced analysis	Community perspectives can improve the interpretation of data and help overcome evaluators' biases and preconceptions.
Builds evaluation capacity in communities	Community participants in the evaluation gain skills that can be applied in future initiatives.
DISADVANTAGES	
Conflicts between academics and communities	Evaluators and community members have very different backgrounds and goals, and these can hamper collaboration.[59]
Time burden	The evaluation process is slower when there is an array of stakeholders involved, and the time burden on community residents can be a deterrent to wide participation.

Advantage/Disadvantage	Rationale
Community engagement skill set deficit	Many evaluators are not trained in community engagement skill sets (a filed unto itself).
Too many voices	Projects may cater to too many voices and are not focused enough to formulate a clear plan or make progress on improving in health outcomes.
Data viewed as less objective	Evaluation may be viewed as less objective because of stakeholder involvement.

The approaches to community engagement in evaluation can be divided into two broad categories: practical participatory evaluation (P-PE) and transformative participatory evaluation (T-PE).[60] P-PE is a more limited approach focusing on the use of evaluation findings to make decisions to improve programs and processes. P-PE is characterized by balanced control of evaluation results between communities and evaluators, modest diversity among stakeholder groups, and an absence of conflict among those with access to different levels of power[61] T-PE is more aligned with Participatory Action Research's focus on power redistribution between communities and researcher/evaluators, but differs from empowerment evaluation[62] in the role of the evaluator because the T-PE evaluator maintains more technical control and is more engaged in managing and directing the evaluation. The "community" that is being engaged is most commonly at the level of program administrator or community representative for P-PE approaches, rather than wide scale lay-person involvement more common in T-PE. There was very limited information about the degree of community engagement in the evaluations we reviewed, only 6 of the 36 mentioned it at all and that was in the context of the intervention rather than the evaluation.

Recommendations

Recommendations in two areas are summarized here: (1) for evaluations of specific initiatives— advice about best practices given the level of available resources; and (2) for the field as a whole— ways of structuring and supporting the gathering of evidence to maximize our ability to draw conclusions about the impact and usefulness of community-level initiatives.

Recommendations for initiative-level evaluation methods

The following are recommendations for evaluation design, data collection and analysis methods, as well as the degree of community engagement in the evaluation. Since evaluation resources are a critical determinant of what can be measured, we provide recommendations for three levels of resources: (1) low, 5%-10% of the intervention budget; (2) medium, 10%-15% of the intervention budget; and (3) high, greater than 15% of the intervention budget. (Note: a typical rule of thumb for evaluation is 10-20% of the intervention budget; and one source recommended between 15-20%).[63]

Evaluation design. The logic model approach is a promising design for community-level initiative evaluations. Given the large number of population-level outcome measures that are typically measured (e.g., a variety of food and physical activity behaviors, obesity rates), some estimate of the population dose or potential impact across outcomes is critical for distinguishing spurious positive findings from those resulting from the initiative and for focusing population-level data collection where effects are most likely to be found. In evaluations with large numbers of communities, regression models can be used to test for associations between environmental changes and population-level outcomes. Analytic

methods for logic model designs should be further developed for attaching statistical significance to observed dose-outcome patterns when there are small numbers of communities.

Measuring environmental changes. Since using a logic model design requires careful measurement of the degree of change that occurs, priority should be given to documenting implementation of environmental change strategies. For modest evaluation budgets, environmental change measurement will likely be limited to progress reporting supplemented by observation of a few key strategies. For more well-resourced evaluations, it should be possible to do pre/post assessments using standard tools and measures. More resources will make it possible to do careful assessments of a greater proportion of the environments that are targeted for intervention.

Strategy-level impact. Another corollary of recommending the logic model approach is that some attempt must be made to estimate the likely impact of the environmental change strategies, in addition to documenting the extent to which the changes take place. This may be done using the population dose framework or other methods of estimating reach and strength. Given the lack of evidence currently in the literature, estimates require opportunistic collection of all the strategy-level data that may be useful in estimating potential impact. Depending on the level of evaluation resources, this can range from the collection of secondary information (e.g., using school cafeteria food purchasing information to track increases in fruit and vegetable consumption resulting from a new salad bar), to more formal evaluation designs (e.g., pre/post surveys of grocer store shoppers to determine the extent to which purchasing patterns have changed as a result of store environmental changes).

Longer-term population-level outcomes. Since the ultimate goal of community-level initiatives is to improve health at the population level, every effort must be made to create ways of tracking these changes over time. For evaluations with limited resources this may mean using secondary data that are often imperfect, either because the measures are not precisely capturing the intended outcomes or because the sample size and geographic focus does not match the intervention geographic boundaries. Primary data collection should focus on the highest dose strategies (i.e., those interventions with the highest reach and strength); rather than data collection such as follow-up surveys of youth and/or adults that may not be needed if all of the intervention strategies targeting the same age groups are low dose. Surveys of youth (e.g., school-based surveys) should take priority when resources are limited: surveying is more cost-effective in schools (versus mail or phone surveys of adults) and most of the initiatives documented as successful in the literature have targeted school-age children.

Table 6 summarizes the above recommendations about evaluation methods for critical evaluation goals: documenting environmental changes, estimating the impact of those changes, and assessing longer-term population-level impact.

Table 6. Recommended Approaches for Key Evaluation Areas, by Level of Evaluation Resources

Resources	Documenting environmental change	Estimating strategy-level impact	Measuring population-level impact
Low (5-10%)[a]	• Oral and written progress reporting annually from community coordinators • Observation of selected key strategies	• Reach and strength estimates based on progress report information and the literature where available	• Secondary data, when available at an appropriate geographic level
Medium (10-15%)	• Oral and written progress reporting at regular intervals jointly by evaluators and community coordinators • Use of environmental and policy assessment tools for selected key strategies	• Reach and strength estimates based on progress reporting information, literature when available, and program evaluations of *selected* key strategies	• Secondary data, if available • School-based surveys of youth food and physical activity attitudes and behaviors
High (>15%)	• Oral and written progress reporting at regular intervals jointly by evaluators and community coordinators • Use of comprehensive and validated environmental and policy assessment tools for all key strategies	• Reach and strength estimates based on progress reporting information, literature when available, and program evaluations of *all* • key strategies	• Secondary data, if available • School-based surveys of youth • Mail/phone surveys of adults

Note:

[a] Percentages indicate the amount of resources for evaluation, as a percent of the intervention budget. .

Community engagement, formative role of evaluation. Community engagement and formative evaluation are critically linked. Without community engagement, there may be inadequate trust in the evaluation process to make strategy improvements based on evaluation findings and recommendations. And an emphasis on providing timely and useful evaluation results for program improvement will aid in getting community buy-in and engagement for the evaluation.

Practical participatory engagement is recommended for evaluations of any resource level, involving key community representatives and organizations (versus community residents alone) in important decisions around evaluation design, data collection instruments and approaches, and analysis/interpretation. The engagement must be "practical," because the time burden is too great for both evaluators and community members for more intensive engagement in evaluations. At the same time, evaluators positioned as part of a team with community members can increase the evaluation capacity of both individuals and organizations.

The formative role of evaluation is enhanced significantly by the use of logic model designs. For example, the population dose concept can provide a useful framework for thinking about ways to increase the impact of obesity prevention strategies,[30] and Fawcett and colleagues have long integrated community planning and implementation with evaluation through the Community Tool Box, a tracking resource for building healthy communities.[64]

Recommendations for advancing the field

This section offers some reflections and recommendations on issues to be addressed to help advance the field of community-based obesity prevention.

Building the evidence base: strategy-level impact. As noted earlier in the discussion about measuring dose, having an estimate of the range of impact for particular strategies (e.g., healthy corner stores, school cafeteria interventions) is critical for both selecting and planning strategies and estimating their potential impact. As we note in an earlier paper,[30] our recommendation is to address this lack of information by encouraging the use of strategy-level evaluations whenever possible and reporting those results in the scientific literature or through web portals such as the AHRQ *Innovations Exchange* (http://www.innovations.ahrq.gov/) or the CDC Division of Community Health's *Success Stories Library* (http://apps.nccd.cdc.gov/dchsuccessstories/searchstories.aspx). The Institute of Medicine recently completed a framework to inform decision making related to the lack of evidence in obesity prevention,[65] and we strongly endorse all of the reports action items related to generating evidence (summary, p. S-9–S-10):

1. Take full advantage of opportunities to generate evidence from ongoing policy and practice.
2. If obesity prevention actions are taken when the evidence is very limited, evaluate the success of the intervention and build credible evidence for use in future decision making.
3. Treat natural experiments, emerging innovations, and ongoing programs as potential sources of useful evidence.
4. Consider forms of evidence and research designs from a variety of disciplines, including systems approaches that can handle complexity.
5. Explore research designs that can be used as alternatives to randomized experiments and that may be more feasible in relation to complex environmental and policy interventions.
6. When reporting results of obesity prevention efforts, include useful aspects of the research related to its generalizability to individuals, settings, contexts, and time frames.

The last two recommendations—employing alternative research designs and reporting research details related to generalizability—are particularly important for generating strategy-level impact estimates. Policy and environmental change strategies are very difficult to evaluate using experimental designs and information from multiple studies using weaker designs may be more productive than attempting a single, large scale experimental study. Reporting details about community context and the way strategies were operationalized can help communities choose and replicate promising approaches most relevant to their own situation.

Building the evidence base: initiative-level impact. Regarding the accumulation of evidence about initiative-level impact, as noted previously, large-scale experimental studies are not feasible; therefore, evidence must be accumulated over time through the evaluation and reporting of a number of individual initiatives. Useful synthesis of this information requires that evaluation reports include:

- community context (size, demographics)
- detailed description of the intervention with facts about the actual implementation (as opposed to planned or intended only)
- assessment of environmental changes that occurred
- population-level results (if any)

Some information about challenges and lessons learned also is valuable.

The accumulation of initiative-level evidence requires that negative studies be published, and it is worth a brief discussion of the problem of publication bias. Publication bias is a serious and increasingly recognized problem across all scientific disciplines.[66, 67] In the field of community-based obesity prevention research, publication bias is even more pronounced. Reviews consistently find almost all positive studies in the peer-reviewed literature—e.g., a recent review of healthy corner store interventions found 9 out of 10 with positive results[68] and our own review of the community-initiative literature found 13 of 16 positive studies—and it is highly unlikely that this represents the true distribution of intervention effectiveness. Several studies have demonstrated the bias against publishing non-significant or negative results.[69, 70] Some of this bias is likely due to self-selection by researcher/evaluators; for example, if a community-level initiative evaluation shows negative findings, there is little incentive for either the funder or evaluator to go to the trouble of publishing that in the literature.

There are additional political factors discouraging the publication of negative findings in the obesity prevention field. There is intense competition for funding at the federal level, combined with an ideological divide related to the degree to which behavior is shaped by individual responsibility or environmental influences. Both sides may use evidence selectively to argue their case so that even one negative study showing that environmental influences were not effective can have a disproportionate impact on policy decisions. This can lead to many researchers who believe in the general principle that environmental influences matter, to self-censor.

It is difficult to have an open and honest debate about which approaches are most effective when negative findings are not published. However, given the barriers to reporting negative studies, encouraging journals to consider negative studies of equal value to positive ones when making publication decisions is important. Another solution may be to increase the availability and inclusion of reports of natural experiments and interim reports of ongoing interventions to fill the gap and avoid the limitations of the available evidence.

Developing and encouraging the use common tools. The aim to have widely used, common tools for both intermediate (e.g., environment, policy change) and longer-term (e.g., food and physical activity behaviors) outcomes is very appealing. A standard set of tools makes it much easier to aggregate results across communities and strategies and make comparisons across initiatives. But there are significant offsetting advantages from tailoring tools to specific initiative goals and study designs. Creating a set of standardized and comprehensive environmental and behavioral assessment tools (e.g., a single environmental assessment for the school environment) may not be the best approach. Instead, we encourage the use of common *measures* (e.g., standard assessment questions for vending machine

contents) where appropriate. Focusing on higher level tool harmonization efforts and creating libraries of standard *measures* rather than standard *tools*, emphasizes the use of common metrics without diluting the utility of the measurement to detect changes in unique settings from new and innovative intervention approaches

However, there may be some cases, particularly related to population-level outcome measures, where it would be useful to encourage the systematic deployment of *tools* that are capable of generating localized results and are already in widespread use. For example, Fitnessgram testing in all schools conducted at a regular interval and grade level, every other year among all fifth graders, would greatly enhance the understanding of prevention efforts across sites.

Summary

In summary, we believe that our recommended approach to evaluating community-level initiatives maximizes useful information about the impact on population-level outcomes for a given level of evaluation resources. Gathering relatively rich information about the actual characteristics of the intervention and the intermediate outcomes such as food and activity behaviors or environmental changes also is preferred because it contributes to our formative understanding of how changes can best be achieved and sustained. The evaluation challenges are substantial but the investment worthwhile given the scale of the investment being made in prevention initiatives to address the nation's obesity epidemic.

Appendix A – Community-Level Obesity Prevention Initiatives– Description and Results

Table A-1. Description of Community-Level Obesity Prevention Initiatives with Population-level Results (n=16)

Initiative	Target population/Design	Intervention	Evaluation methods	Results
Allegiance Health - Health Improvement Organization[71] **(2000-)**[a]	**Setting:** Jackson, MI (Community, USA) **Target population:** Adults, children **Design:** Pre/post	Health partnership efforts among patients, physicians, employers, the health system and the health plan.	Pilot evaluation of worksite wellness component; tracking of employee participation health status measures	**Positive.** Participants managed stress better, avoided weight gain, controlled blood pressure and cholesterol, avoided sick days, and reduced overall health risk.
Arkansas Obesity Prevention Initiative[72] **(2000-2010)**	**Setting:** Arkansas (State-level, USA) **Target population:** Children/adolescents **Design:** Quasi-experimental	Range of statewide efforts to support local schools in making policy and environmental change, including Coordinated School Health and Safe Routes to School grants	School district surveys, stakeholder interviews with parents and school leaders, BMI[b] monitoring. Sample of 484 schools across the state.	**Negative.** No change in obesity rates. Decreases in student purchases from vending machines; but no changes in soda consumption or visits to fast food restaurants.
EPODE[73] **(1992-2004)**	**Setting:** 2 small towns in northern France (Community, Europe) **Target population:** Children, 5-12 years **Design:** Quasi-experimental (post only comparison)	A school-based nutrition information programme initiated in 1992 followed by a number of community-based interventions.	Repeated, cross-sectional, school-based survey for selected school years from 1992-2004 plus BMI on all 5- to 12-year-old children attending school. Survey in comparison towns in 2004 only.	**Positive.** Age-adjusted odds ratio for overweight significantly lower in 2003 and 2004 (girls only). In 2004, the overweight prevalence was significantly lower than in the comparison towns.
5-2-1-0 Let's Go![74] **(2009-2011)**	**Setting:** Greater Portland, ME (Community, USA) **Target population:** Children/adolescents **Design:** Pre/post	Community-level environmental and messaging strategies targeting physical activity, fruits and vegetables, sugary drinks, screentime	Parent surveys from 2007-2011 reporting program awareness and proxy report of children's behavior.	**Positive.** Increased prevalence of targeted behaviors based on parent self- reported data.
Five-a-Day[75] **(2001-2005)**	**Setting:** Five economically deprived communities in England (Community, Europe) **Target population:** Adults **Design:** Quasi-experimental	Community-based interventions to improve fruit and vegetable intake.	There were 975 people in pilot intervention communities compared with 309 people participating in an unrelated observational study as controls.	**Negative.** Increased knowledge and access to fruits and vegetables but no demonstrable effect on total fruit and vegetable intake.

Initiative	Target population/Design	Intervention	Evaluation methods	Results
GEMS (Girls Health Enrichment Multi-site Studies (GEMS)[76, 77] (1999-2001)	**Setting:** Oakland, CA; Memphis, TN (Community, USA) **Target population:** Preadolescent African American girls who were overweight/obese **Design:** Randomized-control trial (individual-level)	Culturally appropriate obesity prevention approaches involving both girls and their parents, community centers or YWCAs, and schools.	Randomized to obesity prevention program intervention or alternative self-esteem building program.	**Negative.** Memphis: no change in BMI. Oakland: changes in BMI were not different in the intervention versus the control group.
Hartslag Limburg[78] (1998-2003)	**Setting:** Maastricht region, Netherlands (Community, Europe) **Target population:** Adults **Design:** Quasi-experimental	Integrative community-based cardiovascular disease prevention program promoting a healthy lifestyle.	Cohort study comparing 5-year mean change in risk factors between the intervention and reference area.	**Positive.** Adjusted difference in mean change in risk factors between intervention and reference group was significant for BMI, waist circumference, total cholesterol and serum glucose.
HEAC/CCROPP (Healthy Eating, Active Communities and Central California Regional Obesity Prevention Program)[79] (2007-2010)	**Setting:** 14 low-income communities in CA (Community, USA) **Target population:** Youth and adults **Design:** Quasi-experimental	Policy and environmental interventions in schools, worksites, health care organizations and the community at large.	Repeated cross sectional surveys of 400 randomly selected 7th and 9th grade students from 13 HEAC communities and 6 out of area comparison communities.	**Positive.** Findings from the school survey combined with environmental assessments confirm that when students are exposed to healthier environments they are more likely to make healthier choices.
HEALTHY Armstrong (Healthy Eating Active Lifestyles Together Helping Youth) [80] (2005-2009)	**Setting:** Rural Armstrong County, PA (Community, USA) **Target population:** Children **Design:** Pre/post	Using elements of the national We Can! program to help children improve their nutritional habits and engage in more physical activity.	Pre- and post-implementation comparisons of student behaviors, including time engaged in physical activity, purchases of high-calorie foods, and school cafeteria expenditures on fresh fruits and vegetables.	**Positive.** Significantly increased levels of physical activity and improved food choices made by students, who consume less "junk food" and more fruits and vegetables in school.
Healthy Hawks Program[81] (2006-)	**Setting:** Communities in Kansas (Community, USA) **Target population:** Overweight children **Design:** Pre/post (individual-level)	Working with children and their family to develop goals and strategies and establish a healthier lifestyle. Community support built for recruitment, and sustainability of changes.	Pre/post BMI; caloric intake (self-reported dietary data).	**Positive.** Significantly reduced caloric intake and BMI among participants after 12 weeks.

Initiative	Target population/Design	Intervention	Evaluation methods	Results
Healthy Living Cambridge Kids[82] **(2004-2007)**	**Setting:** Cambridge, MA (Community, USA) **Target population:** Students K-8 **Design:** Pre/post	Community-based effort in support of the "5-2-1" guidelines: 5+ servings of fruits and vegetables, screen time <2 hours, 1+ hour of exercise.	Comparison of body mass index and fitness test results in a group of 1,900 students tested at baseline and then again 3 years after program implementation.	**Positive.** BMI z-scores and proportion obese decreased, and mean number of fitness tests (0–5) passed increased. Obesity among all race/ethnicity groups declined.
Kaiser Permanente HEAL-CHI (Healthy Eating Active Living Community Health Initiative)[50] **(2006-2010)**	**Setting:** three low-income communities in Northern CA (Community, USA) **Target population:** Youth and adults **Design:** Quasi-experimental logic model design	Policy and environmental interventions in schools, worksites, health care organizations and the community at large.	School-based surveys and Fitnessgram measures of students in intervention and matched comparison communities; surveys of adults using Interactive Voice Response in intervention communities.	**Positive.** Improvements in physical activity behaviors found where high-dose interventions were present in schools.
Nemours Delaware Initiative[83] **(2006-)**	**Setting:** Delaware (State-level, USA) **Target population:** Children **Design:** Quasi-experimental	Statewide policy change, learning collaboratives, technical assistance to schools, childcare and primary care.	Statewide survey in 2006, 2008. Fitnessgram measurement in pilot school PE program (n=19).	**Positive.** Leveling off of obesity rates statewide. Pilot physical education (PE) program in schools showed students in pilot schools 1.5 times more likely to be in healthy fitness zone.
NYC Dept of Health obesity prevention initiative[84] **(2002-)**	**Setting:** New York, NY (Community, USA) **Target population:** Students K-8 **Design:** Pre/post	Community-based environment and policy change efforts, including schools, restaurants, grocery stores, hospitals, worksites.	Use of existing surveys: NYC Community Health Survey, Youth Risk Behavior Survey, NYC Fitnessgram.	**Positive.** Decline in K-8 obesity rate 5.5% between 2006-07 (21.9%) & 2010-11(20.7%), although adult obesity rates increased 18.2% 2002 to 23.4% in 2010.
Romp & Chomp[85] **(2004-2008)**	**Setting:** Geelong, Australia (Community, Australia) **Target population:** Students **Design:** Quasi-experimental	Community-wide, multisetting, multistrategy intervention focused on community capacity building and environmental changes.	Repeat cross-sectional design with a comparison sample.	**Positive.** Significantly lower mean weight, BMI, and BMI z scores in the intervention group. Also significantly lower relative intake of packaged snacks and fruit juice.
Shape Up Somerville[45] **(2002-2005)**	**Setting: Somerville, MA** (Community, USA) **Target population:** Children grades 1-3 **Design:** Quasi-experimental	Comprehensive community-level intervention, involving children, parents, teachers, schools, city departments, healthcare providers.	Non-randomized control trial: 3 intervention compared to 2 comparison schools. Pre/post BMI was primary outcome measure.	**Positive.** BMI z-scores decreased by -0.1005 compared with children in the control communities after controlling for covariates.

Note:

[a] Dates are approximate—often not explicitly included in articles or reports, and sometimes unclear if an initiative is ongoing.

[b] BMI. Body Mass Index is a number calculated from a person's weight and height. BMI provides a reliable indicator of body fatness for most people and is used to screen for weight categories that may lead to health problems.

Table A-2. Description of Community-Level Obesity Prevention Initiatives: In Progress or No Population-level Measurement (n=20)

Initiative	Description	Evaluation methods
CAN DO Houston[86]	Coalition around obesity formed in 2005 – led by a workplace oriented wellness organization. Two pilot neighborhoods selected. Children aged 6-12 years targeted. Focus group approach identified physical activity in one neighborhood (safety) and nutrition education in another.	No outcome evaluation collected – tabulated attendance at after-school programs. Some baseline Fitnessgram data collected.
CLOCC (Chicago)[87]	Obesity prevention coalition in Chicago, $1.9 million budget, variety of activities, programs, collaborations.	Evaluation planned; no details
Collaborate for Healthy Weight[88]	National project of the National Initiative for Children's Healthcare Quality (NICHQ) and the Health Resources and Services Administration (HRSA) bringing together primary care providers, public health professionals, and leaders of community organizations to work across traditional professional borders to address obesity at the community level.	No evaluation details
Communities Putting Prevention to Work (CPPW)[25]	Fifty communities funded (39 obesity prevention) through a 2-year cooperative agreement to reduce chronic disease related to obesity and tobacco using the evidence and practice-based MAPPS[a]. This effort is expected to produce broad, high-impact, sustainable, health outcomes through policy, systems, and environmental change.	National and local evaluations being conducted
Eat Smart, Move More North Carolina[89]	A statewide movement that promotes increased opportunities for healthy eating and physical activity wherever people live, learn, earn, play and pray. Emphasizes policy and organizational change and evidence-based practices (e.g., media campaigns, worksite interventions, BMI monitoring).	Measures selected for tracking: consumption of sugar-sweetened beverages and increased opportunities for extracurricular activity
Get a Life! (Mississippi)[90]	Supports schools, churches, local governments, and employers in eight rural Mississippi counties in addressing the area's obesity epidemic. Key program elements include: supporting local health councils, providing technical support, and regional planning.	Tracking implementation of activities of various stakeholders targeted by the program; some anecdotal reports of improvements in health-related behaviors.
Go for Your Life (Victoria, Australia)[91]	Community-based interventions in 6 communities in regions of low socioeconomic status. Planned and managed by primary care physicians (PCPs)/lead agencies, support from Department of Health Services DHS and a state-wide evaluator.	Evaluation being conducted with control communities, repeat cross-sectional measures of impact/outcome.
Healthy Alberta Communities Project (Alberta Provence, Canada)[92]	Partnership between the Health Ministry and university of Alberta to promote environmental approaches to obesity prevention.	Described and reflected critically upon the level and nature of community capacity built

Initiative	Description	Evaluation methods
Healthy and Active Communities (Missouri)[93]	Approaches include grantmaking, evaluation support, dissemination technical assistance, policy assessment, and development of local, regional, and statewide collaborations to increase access to physical activity and nutrition through environmental, policy and behavior change.	External evaluators conducting an overall longitudinal evaluation; technical assistance being provided for local evaluations. Baseline policy assessment conducted.
Healthy Communities Study[44]	Five-year observational study of communities that aims to (1) determine the associations between community programs/policies and body mass index (BMI), diet, and physical activity in children; and (2) identify the community, family, and child factors that modify or mediate the associations between community programs/policies and BMI, diet, and physical activity in children.	A cross-sectional assessment (e.g., survey, medical chart abstraction of heights and weights for up to 10 years prior to in-person measurement, observation of school nutrition and physical activity) of 268 communities and over 21,000 children in grades K-8 and their parents, along with a detailed review of policies/programs in place in the communities.
Healthy Eating Active Living Cities Campaign (California)[94]	Builds awareness among California city officials about the role of the physical environment in promoting healthy habits, and provides them with an array of practical support for passing policies and resolutions to make it easier for residents to engage in healthy behaviors.	Post-implementation data available on the number of cities that have adopted policies and resolutions to support behavior change among residents.
Healthy Kids, Healthy Communities[23]	Nationwide initiative in 50 communities pursuing policy & environmental change strategies.	Progress reporting only.
IDEFICS (Identification and prevention of dietary- and lifestyle-induced health effects in children and infants)[95]	Developed and implemented innovative community-oriented intervention programmes for obesity prevention and healthy lifestyle primarily in children aged 2-10 years in eight European countries: Sweden, Estonia, Germany, Belgium, Hungary, Italy, Spain and Cyprus. Eight matched pair communities per country.	Long-term outcome design with 64 communities. The overall intervention programme's duration was 2 years, but a longer-term follow-up programme is being developed.
Project FIT (Grand Rapids, MI)[96]	Collaboration between the public school system, local health systems, physicians, neighborhood associations, businesses, faith-based leaders, community agencies and university researchers to develop a multi-faceted approach to promote physical activity and healthy eating.	Pre-post survey evaluation.
Recreation Rx (San Diego, CA)[97]	Facilitates partnerships between physicians and recreation providers in underserved communities to increase access to safe and structured activities.	Pre/post physician surveys, program utilization statistics, anecdotal reports.
San Diego County Childhood Obesity Initiative[98]	Public/private partnership to reduce and prevent childhood obesity in San Diego County by creating healthy environments for all children and families through advocacy, education, policy development, and environmental change.	No results yet. Quasi-experimental design using paper baseline- and post-surveys of youth, on-line surveys of adult mentors, and interviews with decision-makers who are the focus of advocacy efforts.

Initiative	Description	Evaluation methods
Shape NC (Smart Start & The North Carolina Partnership for Children)[99]	Creating a cadre of early childhood health and wellness champions among state and local leaders and the professionals working with young children and families, and; ensuring that children attending child care programs are served nutritious foods, engage in physical activity, and have teachers modeling healthy behaviors.	Nutrition and Physical Activity Self-Assessment for Child Care (NAP SACC) survey (pre/post).
Wayne County Health Department/Partnership for the Children of Wayne County (NC)[100]	Partnership working with nonprofit groups to promote better nutrition and increased physical activity among preschoolers who attend 8 local childcare centers.	Pre/post implementation and comparisons of practices related to nutrition and physical activity at participating child care centers; post survey of parents' behaviors and perceptions.
WE CAN![101]	National movement that offers organizations, community groups, and health professionals a centralized resource to promote a healthy weight in youth through community outreach, partnership development, and media activities.	Fourteen Intensive Community Sites were selected to implement WE CAN! programming for at least one year. Pre/post surveys related to parent and student curricula are planned.
WK Kellogg Foundation Food and Fitness Initiative[22]	Creating communities that support access to locally grown, healthy, affordable food, and safe and convenient places for physical activity and play, for families and children. Nine communities nationwide funded for implementation.	Evaluations in progress.

Note:

[a] MAPPS. Five evidence-based strategies, when combined, expected to improve health behaviors by changing community environments: Media, Access, Point of decision information, Price, and Social support/services.

References

1. Ebbeling CB, Pawlak DB, Ludwig DS. Childhood obesity: public-health crisis, common sense cure. *Lancet.* Aug 10 2002;360(9331):473-482.

2. Kumanyika SK, Obarzanek E, Stettler N, et al. Population-based prevention of obesity: the need for comprehensive promotion of healthful eating, physical activity, and energy balance: a scientific statement from American Heart Association Council on Epidemiology and Prevention, Interdisciplinary Committee for Prevention (formerly the expert panel on population and prevention science). *Circulation.* Jul 22 2008;118(4):428-464.

3. U.S. Office of the Surgeon General. *The Surgeon General's Call to Action to Prevent and Decrease Overweight and Obesity*. Rockville, MD: U.S. Dept. of Health and Human Services; 2001.

4. Woodward-Lopez G. *Obesity : Dietary and Developmental Influences*. Boca Raton, FL: CRC/Taylor & Francis; 2006.

5. Cohen DA, Scribner RA, Farley TA. A structural model of health behavior: a pragmatic approach to explain and influence health behaviors at the population level. *Prev Med.* Feb 2000;30(2):146-154.

6. McLeroy KR, Bibeau D, Steckler A, Glanz K. An ecological perspective on health promotion programs. *Health Educ Q.* Winter 1988;15(4):351-377.

7. COMMIT Research Group. Community Intervention Trial for Smoking Cessation (COMMIT): I. cohort results from a four-year community intervention. *Am J Public Health.* Feb 1995;85(2):183-192.

8. Farquhar JW, Fortmann SP, Flora JA, et al. Effects of communitywide education on cardiovascular disease risk factors. The Stanford Five-City Project. *JAMA.* Jul 18 1990;264(3):359-365.

9. Wagner EH, Wickizer TM, Cheadle A, et al. The Kaiser Family Foundation Community Health Promotion Grants Program: findings from an outcome evaluation. *Health Serv Res.* Aug 2000;35(3):561-589.

10. Merzel C, D'Afflitti J. Reconsidering community-based health promotion: promise, performance, and potential. *Am J Public Health.* Apr 2003;93(4):557-574.

11. Brownson RC, Haire-Joshu D, Luke DA. Shaping the context of health: a review of environmental and policy approaches in the prevention of chronic diseases. *Annu Rev Public Health.* 2006;27:341-370.

12. Pruss-Ustun A, Corvalan C. How much disease burden can be prevented by environmental interventions? *Epidemiology.* Jan 2007;18(1):167-178.

13. Ashe M, Feldstein LM, Graff S, Kline R, Pinkas D, Zellers L. Local venues for change: legal strategies for healthy environments. *J Law Med Ethics.* Spring 2007;35(1):138-147.

14. Doyle S, Kelly-Schwartz A, Schlossberg M, Stockard J. Active community environments and health: The relationship of walkable and safe communities to individual health. *J Am Plann Assoc.* 2012/12/26 2006;72(1):19-31.

15. Glanz K, Hoelscher D. Increasing fruit and vegetable intake by changing environments, policy and pricing: restaurant-based research, strategies, and recommendations. *Prev Med.* Sep 2004;39 Suppl 2:S88-93.

16. Sallis JF, Cervero RB, Ascher W, Henderson KA, Kraft MK, Kerr J. An ecological approach to creating active living communities. *Annu Rev Public Health.* 2006;27:297-322.

17. Sallis JF, Glanz K. The role of built environments in physical activity, eating, and obesity in childhood. *Future Child.* Spring 2006;16(1):89-108.

18. Khan LK, Sobush K, Keener D, et al. Recommended community strategies and measurements to prevent obesity in the United States. *MMWR Recomm Rep.* Jul 24 2009;58(RR-7):1-26.

19. Koplan J, Liverman CT, Kraak VI, IOM Committee on Prevention of Obesity in Children and Youth. *Preventing Childhood Obesity: Health in the Balance*. Washington: National Academies Press; 2005.

20. Parker L, Burns AC, Sanchez E, IOM Committee on Childhood Obesity Prevention Actions for Local Governments. *Local Government Actions to Prevent Childhood Obesity*. Washington, DC: National Academies Press; 2009.

21. Story M, Kaphingst KM, French S. The role of schools in obesity prevention. *Future Child.* Spring 2006;16(1):109-142.

22. W.K. Kellogg Foundation. Food and Fitness Initiative. http://www.nifa.usda.gov/nea/food/in_focus/health_if_food_fitness_initiative.html.

23. Robert Wood Johnson Foundation. Healthy Kids, Healthy Communities. http://www.healthykidshealthycommunities.org/.

24. Cheadle A, Schwartz PM, Rauzon S, Beery WL, Gee S, Solomon L. The Kaiser Permanente Community Health Initiative: overview and evaluation design. *Am J Public Health.* Nov 2010;100(11):2111-2113.

25. Centers for Disease Control and Prevention. Communities Putting Prevention to Work. http://www.cdc.gov/communitiesputtingpreventiontowork/.

26. Let's Move! http://www.letsmove.gov/.

27. The White House Office of the Press Secretary. Presidential Memorandum -- Establishing a Task Force on Childhood Obesity. 2010; http://www.whitehouse.gov/the-press-office/presidential-memorandum-establishing-a-task-force-childhood-obesity. Accessed April 20, 2010.

28. Centers for Disease Control and Prevention. Community Transformation Grants (CTG). http://www.cdc.gov/communitytransformation/index.htm.

29. Fawcett SB, Schultz J. Supporting participatory evaluating using The Community Toolbox online documentation system. In: Minkler M, Wallerstein N, eds. *Community-based Participatory Research for Health : From Process to Outcomes*. 2nd ed. San Francisco: Jossey-Bass; 2008.

30. Cheadle A, Schwartz PM, Rauzon S, Bourcier E, Senter S, Spring R. Using the concept of "population dose" in planning and evaluating community-level obesity prevention initiatives. *Am J Eval.* November 6, 2012 2012.

31. Community Preventive Services Task Force. The Guide to Community Preventive Services: What Works to Promote Health. http://www.thecommunityguide.org/index.html.

32. University of Wisconsin Population Health Institute. What Works for Health. http://www.countyhealthrankings.org/programs.

33. Bodor JN, Rose D, Farley TA, Swalm C, Scott SK. Neighbourhood fruit and vegetable availability and consumption: the role of small food stores in an urban environment. *Public Health Nutr.* Apr 2008;11(4):413-420.

34. Rundle A, Neckerman KM, Freeman L, et al. Neighborhood food environment and walkability predict obesity in New York City. *Environ Health Perspect.* Mar 2009;117(3):442-447.

35. Booth KM, Pinkston MM, Poston WS. Obesity and the built environment. *J Am Diet Assoc.* May 2005;105(5 Suppl 1):S110-117.

36. Heath GW, Brownson RC, Kruger J, et al. The effectiveness of urban design and land use and transport policies and practices to increase physical activity: A systematic review. *J Phys Act Health.* 2006;3(Suppl 1):S55-S76.

37. Papas MA, Alberg AJ, Ewing R, Helzlsouer KJ, Gary TL, Klassen AC. The built environment and obesity. *Epidemiol Rev.* 2007;29:129-143.

38. Berkson SS, Espinola J, Corso KA, Cabral H, McGowan R, Chomitz VR. Reliability of Height and Weight Measurements Collected by Physical Education Teachers for a School-Based Body Mass Index Surveillance and Screening System. *J Sch Health.* Jan 2013;83(1):21-27.

39. Homer J, Milstein B, Wile K, et al. Simulating and evaluating local interventions to improve cardiovascular health. *Prev Chronic Dis.* Jan 2010;7(1):A18.

40. Koepsell TD, Wagner EH, Cheadle AC, et al. Selected methodological issues in evaluating community-based health promotion and disease prevention programs. *Annu Rev Public Health.* 1992;13:31-57.

41. Atienza AA, King AC. Community-based health intervention trials: an overview of methodological issues. *Epidemiol Rev.* 2002;24(1):72-79.

42. Bunin GR, Spector LG, Olshan AF, et al. Secular trends in response rates for controls selected by random digit dialing in childhood cancer studies: a report from the Children's Oncology Group. *Am J Epidemiol.* Jul 1 2007;166(1):109-116.

43. Curtin R, Presser S, Singer E. Changes in Telephone Survey Nonresponse over the Past Quarter Century. *Public Opin Q.* Spring 2005 2005;69(1):87-98.

44. National Heart L, and Blood Institute. Healthy Communities Study. http://www.nhlbi.nih.gov/hcs/index.htm.

45. Economos CD, Hyatt RR, Goldberg JP, et al. A community intervention reduces BMI z-score in children: Shape Up Somerville first year results. *Obesity (Silver Spring).* May 2007;15(5):1325-1336.

46. Cheadle A, Beery WL, Greenwald HP, Nelson GD, Pearson D, Senter S. Evaluating the California Wellness Foundation's Health Improvement Initiative: a logic model approach. *Health Promot Pract.* Apr 2003;4(2):146-156.

47. Francisco VT, Paine AL, Fawcett SB. A methodology for monitoring and evaluating community health coalitions. *Health Educ Res.* Sep 1993;8(3):403-416.

48. Collie-Akers VL, Fawcett SB, Schultz JA, Carson V, Cyprus J, Pierle JE. Analyzing a community-based coalition's efforts to reduce health disparities and the risk for chronic disease in Kansas City, Missouri. *Prev Chronic Dis.* Jul 2007;4(3):A66.

49. Glasgow RE, Klesges LM, Dzewaltowski DA, Estabrooks PA, Vogt TM. Evaluating the impact of health promotion programs: using the RE-AIM framework to form summary measures for decision making involving complex issues. *Health Educ Res.* Oct 2006;21(5):688-694.

50. Cheadle A, Rauzon S, Spring R, et al. Kaiser Permanente's Community Health Initiative in Northern California: evaluation findings and lessons learned. *Am J Health Promot.* Nov-Dec 2012;27(2):e59-68.

51. Saelens BE, Glanz K. Work group I: Measures of the food and physical activity environment: instruments. *Am J Prev Med.* Apr 2009;36(4 Suppl):S166-170.

52. Murray DM. *Design and analysis of group-randomized trials*. New York: Oxford University Press; 1998.

53. Bender R, Lange S. Adjusting for multiple testing--when and how? *J Clin Epidemiol.* Apr 2001;54(4):343-349.

54. Hochberg Y, Tamhane AC. *Multiple comparison procedures*. New York: Wiley; 1987.

55. Minkler M, Wallerstein N. *Community-based Participatory Research for Health : From Process to Outcomes*. 2nd ed. San Francisco: Jossey-Bass; 2008.

56. CDC ASTDR Committee on Community Engagement. *Principles of Community Engagement*. Atlanta, GA: Centers for Disease Control and Prevention, Public Health Practice Program Office; 1997.

57. Krishnaswami J, Martinson M, Wakimoto P, Anglemeyer A. Community-engaged interventions on diet, activity, and weight outcomes in U.S. schools: a systematic review. *Am J Prev Med.* Jul 2012;43(1):81-91.

58. Cousins JB, Chouinard JA. *Participatory Evaluation Up Close : An Integration of Research-based Knowledge*. Charlotte, N.C.: Information Age Publishing; 2012.

59. Economos CD, Tovar A. Promoting health at the community level: thinking globally, acting locally. *Child Obes.* Feb 2012;8(1):19-22.

60. King JA, Cousins JB, Whitmore E. Making sense of participatory evaluation: Framing participatory evaluation. *New Directions for Evaluation.* 2007;2007(114):83-105.

61. Shaw IF, Greene JC, Mark MM. *Handbook of Evaluation*. London: SAGE; 2006.

62. Fetterman DM, Wandersman A. *Empowerment Evaluation Principles in Practice*. New York: Guilford Press; 2005.

63. Bronte-Tinkew J, Joyner K, Allen T. *Five Steps for Selecting an Evaluator: A Guide for Out-of-School Time Practitioners*: Child Trends; 2007.

64. University of Kansas Work Group for Community Health and Development. The Community Tool Box. http://ctb.ku.edu/en/default.aspx.

65. Kumanyika SK, Parker L, Sim LJ, IOM Committee on an Evidence Framework for Obesity Prevention Decision Making. *Bridging the Evidence Gap in Obesity Prevention : A Framework to Inform Decision Making*. Washington, D.C.: National Academies Press; 2010.

66. Ioannidis JP. Why most published research findings are false. *PLoS Med.* Aug 2005;2(8):e124.

67. Rosenthal R. The file drawer problem and tolerance for null results. *Psychol Bull.* 1979;86(3):638-641.

68. Gittelsohn J, Rowan M, Gadhoke P. Interventions in small food stores to change the food environment, improve diet, and reduce risk of chronic disease. *Prev Chronic Dis.* Feb 2012;9:E59.

69. Ioannidis JP. Effect of the statistical significance of results on the time to completion and publication of randomized efficacy trials. *JAMA.* Jan 28 1998;279(4):281-286.

70. Stern JM, Simes RJ. Publication bias: evidence of delayed publication in a cohort study of clinical research projects. *BMJ.* Sep 13 1997;315(7109):640-645.

71. AHRQ Health Care Innovations Exchange. Multistakeholder, community-wide collaborative prevents disease and promotes health. 2012; http://www.innovations.ahrq.gov/content.aspx?id=1755.

72. University of Arkansas for Medical Sciences. *Year 7 Evaluation: Arkansas Act 1220 of 2003 to Combat Childhood Obesity.* Little Rock, AR: Fay W. Boozman College of Public Health; 2011.

73. Romon M, Lommez A, Tafflet M, et al. Downward trends in the prevalence of childhood overweight in the setting of 12-year school- and community-based programmes. *Public Health Nutr.* Oct 2009;12(10):1735-1742.

74. AHRQ Health Care Innovations Exchange. Tools, education, and technical assistance encourage child care sites to adopt strategies that promote health eating and physical activity in young children. 2012; http://www.innovations.ahrq.gov/content.aspx?id=3253.

75. Ashfield-Watt PA, Welch AA, Godward S, Bingham SA. Effect of a pilot community intervention on fruit and vegetable intakes: use of FACET (Five-a-day Community Evaluation Tool). *Public Health Nutr.* Jul 2007;10(7):671-680.

76. Klesges RC, Obarzanek E, Kumanyika S, et al. The Memphis Girls' health Enrichment Multi-site Studies (GEMS): an evaluation of the efficacy of a 2-year obesity prevention program in African American girls. *Arch Pediatr Adolesc Med.* Nov 2010;164(11):1007-1014.

77. Robinson TN, Matheson DM, Kraemer HC, et al. A randomized controlled trial of culturally tailored dance and reducing screen time to prevent weight gain in low-income African American girls: Stanford GEMS. *Arch Pediatr Adolesc Med.* Nov 2010;164(11):995-1004.

78. Schuit AJ, Wendel-Vos GC, Verschuren WM, et al. Effect of 5-year community intervention Hartslag Limburg on cardiovascular risk factors. *Am J Prev Med.* Mar 2006;30(3):237-242.

79. *Healthy Eating, Active Communities and Central California Regional Obesity Prevention Program Final Evaluation Synthesis Report.* Oakland, CA: Samuels & Associates; 2011.

80. AHRQ Health Care Innovations Exchange. Community coalition supports schools in helping students increase physical activity and make better food choices. 2012; http://www.innovations.ahrq.gov/content.aspx?id=3232.

81. AHRQ Health Care Innovations Exchange. Staged weight management program helps overweight children reduce caloric intake and body mass index. 2012; http://www.innovations.ahrq.gov/content.aspx?id=2290.

82. Chomitz VR, McGowan RJ, Wendel JM, et al. Healthy Living Cambridge Kids: a community-based participatory effort to promote healthy weight and fitness. *Obesity (Silver Spring).* Feb 2010;18 Suppl 1:S45-53.

83. Chang DI, Gertel-Rosenberg A, Drayton VL, Schmidt S, Angalet GB. A statewide strategy to battle child obesity in Delaware. *Health Aff (Millwood).* Mar-Apr 2010;29(3):481-490.

84. *Reversing the Epidemic: The New York City Obesity Task Force Plan to Prevent and Control Obesity.* 2012.

85. de Silva-Sanigorski AM, Bell AC, Kremer P, et al. Reducing obesity in early childhood: results from Romp & Chomp, an Australian community-wide intervention program. *Am J Clin Nutr.* Apr 2010;91(4):831-840.

86. Correa NP, Murray NG, Mei CA, et al. CAN DO Houston: a community-based approach to preventing childhood obesity. *Prev Chronic Dis.* Jul 2010;7(4):A88.

87. Becker AB, Longjohn M, Christoffel KK. Taking on childhood obesity in a big city: Consortium to Lower Obesity in Chicago Children (CLOCC). *Progress in Pediatric Cardiology.* 2008;25(2):199-206.

88. National Initiative for Children's Healthcare Quality. *Joining Forces for Healthier Communities: Collaborate for Healthy Weight Progress Report.* 2012.

89. Eat Smart, Move More North Carolina. http://www.eatsmartmovemorenc.com/index.html.

90. AHRQ Health Care Innovations Exchange. Foundation supports rural stakeholders in promoting better eating and physical activity, leading to anecdotal reports of improved behaviors and outcomes. 2012; http://www.innovations.ahrq.gov/content.aspx?id=3267.

91. Haby M, Doherty R, Holt L. *Experiences of the Health Promoting Communities demonstration projects [PowerPoint presentation].* Melbourne, Victoria: Department of Human Services, Health Development Unit, Evidence, Evaluation & Policy Team; 2009.

92. Raine KD, Plotnikoff R, Nykiforuk C, et al. Reflections on community-based population health intervention and evaluation for obesity and chronic disease prevention: the Healthy Alberta Communities project. *Int J Public Health.* Dec 2010;55(6):679-686.

93. Hessel AS, Marshall JW, Brown W, Sabina AB, DeForest K. Healthy & Active Communities Initiative: a foundation's response to the obesity epidemic. *Prev Med.* Jan 2010;50 Suppl 1:S93-94.

94. AHRQ Health Care Innovations Exchange. Statewide campaign provides education and technical support to city officials, leading to policies that promote healthy behaviors. 2012; http://www.innovations.ahrq.gov/content.aspx?id=3231.

95. De Henauw S, Verbestel V, Marild S, et al. The IDEFICS community-oriented intervention programme: a new model for childhood obesity prevention in Europe? *Int J Obes (Lond).* Apr 2011;35 Suppl 1:S16-23.

96. Eisenmann JC, Alaimo K, Pfeiffer K, et al. Project FIT: rationale, design and baseline characteristics of a school- and community-based intervention to address physical activity and healthy eating among low-income elementary school children. *BMC Public Health.* 2011;11:607.

97. AHRQ Health Care Innovations Exchange. Recreation "prescriptions" increase use of free community exercise programs by low-income patients who are overweight or obese. 2012; http://www.innovations.ahrq.gov/content.aspx?id=2934.

98. San Diego County Childhood Obesity Initiative. http://ourcommunityourkids.org/.

99. Iruka I, Maxwell K, Sotolongo J. *Nutition and Physical Activity Self Assessment for Child Care (NAP SACC): Final 2009 Evaluation Report*: University of North Carolina FPG Child Development Institute & North Carolina Partnership for Children, Inc.

100. AHRQ Health Care Innovations Exchange. County, city, and community agencies support childcare centers and parents in improving nutrition and physical activity habits of preschoolers. 2012; http://www.innovations.ahrq.gov/content.aspx?id=3283.

101. National Heart L, and Blood Institute. We Can! Ways to Enhance Children's Activity and Nutrition. http://www.nhlbi.nih.gov/health/public/heart/obesity/wecan/.