Expanding access to high-quality medical care may improve health status, but the biggest advances in health may actually come from addressing upstream determinants of health by creating communities where people can thrive. Sharing long-term goals, public health and community development professionals can advance healthy communities together if health considerations are incorporated into project planning, implementation and evaluation. In this essay, we recommend incorporating evaluation into project planning early on; clearly defining causes and hoped for effects with practical and affordable evaluation designs; prioritizing measurement; standardizing, simplifying, and innovating measurement methods; and maximizing collaboration between public health and community developers.
EVALUATING SOCIAL DETERMINANTS OF HEALTH IN COMMUNITY DEVELOPMENT PROJECTS

David Fleming, Hilary N. Karasz, and Kirsten Wysen
Public Health-Seattle & King County

We know the all-too-familiar statistics—despite the United States spending almost twice as much as other developed countries on health care, we rank 33rd in life expectancy. Although expanding access to high-quality health care is important, improvements in medical care delivery can be expected to reduce preventable deaths by only about 10 percent.\(^1\) To make bigger improvements in health, we need to look outside the walls of the clinic to the places where we live, learn, work, and play. Individual health is based in large part on the health of the communities in which one lives, homes free of toxins, accessible parks and community centers, convenient and safe transit, biking and walking trails, and access to good food, child care, and jobs. These “social” or “upstream” community determinants of health can influence as much as 60 percent of preventable mortality.\(^2\)

Almost without exception, the goals of community development projects include improving one or more of these upstream determinants of health. In theory, then, a logical outcome of successful community development should be improved community health, although this is neither a necessary nor intentional outcome. Also, in theory, evaluations quantifying the health improvements resulting from well-designed community development projects should be plentiful and broadly disseminated. Unfortunately, practice has not caught up with theory,

---

2. Ibid.
and our evaluation cupboards are mostly bare. In part, this may be the result of not prioritizing evaluations, but in truth there are significant obstacles to successfully measuring the effects of community development projects on health.

One early example of this kind of upstream intervention was recent multi-sector work in King County, Washington, to improve school nutrition and student physical activity and reduce obesity in some of the poorest school districts in the county. This work, led by the public health department of Seattle and King County, involved the collaboration of K-12 education, the food system, urban planning, small business, and other sectors to make concentrated investments in specific locations.

At the end of the two-year initiative, results showed that for children living in the project area, obesity prevalence dropped by a highly significant 17 percent, whereas obesity remained unchanged in other parts of the county.3 Hidden behind this one-sentence result are lessons in front-line complexities of conducting this type of evaluation. Drawing in part from this experience, the essay outlines some of the steps that should be considered in community development projects that aim to improve the health of residents, among other goals.

**INCORPORATE EVALUATION INTO PROJECT PLANNING EARLY**

Evaluation should be part of project planning at the earliest conceptualization phase. Too often, evaluation is an add-on after the bulk of planning has been done, when time and budget are short and it is too late to make significant changes to the project.

The King County project had a strong evaluation component from the start. Evaluators played an important role in project design and were active participants during the start-up phase of the project, when adjustments were made to the project design based on early results. Almost 10 percent of the total project cost was dedicated to data collection and evaluation, a signal of the importance of evaluation to the project.

---

Incorporating evaluation early is not only more efficient than doing evaluation post-hoc, it also may allow for more deliberate incorporation of health interventions. If health is a desired goal of a housing development project, planners can think through in advance how to incorporate opportunities for residents to be physically active. They can install sidewalks and bike paths, provide green spaces and walking destinations such as shopping, and take into account transit routes so people can walk to and from buses and trains. This explicit planning simplifies evaluation. If walking paths are being incorporated into a project to increase exercise, it becomes clearer that a primary goal of an evaluation will be to determine whether these paths are being used.

Thinking about evaluation early may also create opportunities for more elegant evaluation through smart project implementation. For example, a larger scale renovation project may allow for testing of the impact of design elements by sequencing the construction and comparing results in renovated versus not yet renovated areas (newer buildings, for example, may prioritize the placement of the stairway to encourage using the stairs instead of elevators, unlike older buildings).

**CLEARLY DEFINE CAUSES AND HOPED-FOR EFFECTS. THEN PICK A PRACTICAL AND AFFORDABLE EVALUATION DESIGN FOR MEASURING THEM**

One of the most difficult evaluation barriers to overcome is the overpowering conventional wisdom that the best (and in some minds the only) way to scientifically conduct a valid evaluation is through a “gold standard” randomized, double blinded, placebo-controlled trial (RCT). RCTs eliminate many sources of bias and are a great design for identifying whether a new drug or vaccine works—but usually not for whether a community development project is improving health. Using RCTs to evaluate the effects of upstream determinants of health is plagued by problems, including the impracticality of random assignment of subjects, the difficulty in limiting exposure of the intervention to just the experimental group, the impossibility of blinding people or researchers to whether they have received the intervention, and the costs.
There are many alternative, more practical and less expensive approaches to evaluation, and community developers have an important co-conspirator here: the public health scientist who spends his or her life in the same “real world” trying to evaluate the same kinds of interventions. Examples of interventions with strong evidence bases from a variety of study designs can be found in the Centers for Disease Control and Prevention’s (CDC) “Community Guide.” This online resource is an important collection of evidence-based practices, which continues to expand.⁴

The bottom line is that an evaluation should help determine whether what was done did what it was hoped it would. At its core, evaluations usually involve measuring what was done, what happened, and, in some way, what would have happened if nothing had been done (the reason for the placebo arm of an RCT). As a consequence, one of the most important features of any evaluation design is a comparison group that did not receive the benefit of the intervention. The use of a comparison group dramatically strengthens the power of the evaluation but most often entails additional data collection costs.

Different types of comparison groups are possible. Quasi-experimental elements such as the 1811 Eastlake study of medical care costs for homeless people with alcoholism, for example, used such a design to compare 95 housed participants (with drinking permitted) with 39 wait-list control participants.⁵ The results showed impressive health improvements and cost savings, and this type of Housing First approach is being adopted throughout the country after years of little agreement on how to reduce chronic homelessness. Other designs are simpler still. Examples include sequential implementation, in which a comparison group receives the intervention after a group that receives the intervention in the first round. Another example is a pre/post design, in which a group serves as its own comparison by measuring changes in the group

---

⁴ The Cochrane Review and the Coalition for Evidence-based Policy also have public access compendiums of a small but growing number of proven practices along with indications of the strength of the evidence about community interventions.

before and after the intervention (yet another reason to incorporate evaluation early).

The problem with a comparison group that has not been randomly drawn from an initial set of eligible participants is that it may not be comparable and observed differences between the groups may be for reasons other than the intervention. Comparing drug treatment outcomes between individuals who voluntarily enter treatment with those who decline, for example, may reflect differences in motivation to quit rather than effectiveness of treatment. Statistical and study design elements can minimize these problems. For example, regression analysis can separate the effects of the chosen intervention from other known influences and can limit the chance of faulty conclusions owing to confounding factors extraneous to the intervention (a type of analysis where evaluators can really earn their pay).

In the King County obesity reduction project, project evaluators used existing data sources to compare obesity data in the intervention districts with similar data in the other 12 King County districts that did not participate. The two groups were somewhat dissimilar in that the intervention districts were lower income and more poorly resourced, while the nonintervention districts were much higher income. The risk
was that natural improvements in the wealthier districts could have obscured intervention group improvement.

Community development evaluations must overcome the challenge assessing cause and effect in which interventions (such as housing or preschool education) commonly act through more than one pathway to improve health and also may influence more than one health outcome (Figure 1).

Logic models—visual depictions of the relationships between inputs and outputs—can help overcome this difficulty by providing a way to see and agree on the interventions, outcomes, and the intervening pathways. For example, Figure 2 shows a simple logic model describing the relationship between bike paths and health.

Actual logic models likely will be more detailed, particularly if there are multiple interventions and outcomes. In community development projects that address health risk factors, logic models can be integrated into the original project design and evaluation plan. Developing a logic model allows project managers to talk through assumptions about how design decisions will impact health and set expectations about short term, medium term, and long term impact.

**PRIORITIZE MEASURING WHAT COMES FIRST**

Sometimes, in evaluating interventions to improve health, it is easy to fall prey to a “tyranny of outcomes” mindset that prioritizes measuring
Adopting More Strategic Practices

important health outcomes above all else. This desire may be intensified by a belief that potential funders will be most receptive to a promise of “hard proof” that their money has improved health. Like the conventional wisdom that the randomized control trial is the gold standard of evaluation design, this “tyranny of outcomes” mindset is often incorrect and can lead to inefficient, ineffective approaches.

By their nature, community development projects are most likely to be targeted at upstream determinants of health (bike paths as in the logic model example above, increased social connectedness, better access to healthy food, etc.). Generally, evaluations should prioritize identifying and measuring the earliest expected outcomes. One advantage in using a well-planned logic model to guide evaluation is that it should indicate what these first outcomes are. Measuring the most important earliest signs of success will show whether the project is on the pathway to better health. If it is not, the program managers may want to revise the intervention or modify the expected outcomes.

For example, an evaluation of a housing project with nutrition and physical activity attributes designed to reduce the rates of diabetes should not focus first on measuring changes in diabetes rates. Instead, evaluation should first monitor whether the planned health assets were implemented on time and as designed (good project management). In this housing project example, the first step may be to ensure that the exercise facility for the housing project was completed and opened. Subsequently, evaluators can measure how frequently it is being used. If time and resources allow, the next step would be to think about short-term indicators of health in residents using it. Finally, for well-funded projects, longer-term follow-up could more clearly establish the evidence base that these short-term outcomes do lead to longer-term health improvements.

Optimally, evaluations should prioritize measuring intermediate outcomes that have been shown by prior research to be connected to longer-term outcomes. For example, antismoking interventions routinely measure changes in tobacco use rather than waiting 30 years to measure reduction in lung cancer mortality. Similarly, an intervention to increase physical activity by installing bike paths probably should not target its
evaluation dollars to measuring reduction in heart disease, but rather look farther upstream in the logic model, to links between exercise and obesity reduction and less hypertension. The most important measure is probably one of bike use. The downstream health outcomes such as causes of death can be left for another day, assuming the interim outcome measures along the way (and the budget) are adequate.

The King County obesity prevention project’s intent was to improve health in the targeted school districts through policy, systems, and environmental change. Its logic models showed that the first significant outcome to measure was the number and type of policies that changed, such as adoption of physical education curriculum or lunchroom cafeteria food purchase standards. Had no policies changed, then measuring health outcomes would have been a wasted effort. The intermediate outcome in this case was obesity prevalence, and the long-term outcome, something that would not be measurable for a decade or more, was heart and other chronic disease where obesity is a risk factor.

Greater use of intermediate outcomes, process measures, and rapid-cycle surveys can speed the process of obtaining information measures sufficient to take action and make mid-course corrections by rolling with interim results rather than waiting for the results of lengthy trials. Micro-trials are becoming more common in medical research, and in these early days of community development/public health collaboration, it makes sense to measure as many project approaches as possible to rapidly determine what works and to eliminate approaches that do not. New research methods, such as rapid online survey data collection and text responses, can provide nearly real-time feedback.

**STANDARDIZE, SIMPLIFY, AND BE INNOVATIVE ABOUT MEASUREMENT METHODS AND APPROACHES**

In addition to building a database of what works, the pathway to more effective and efficient evaluations will be aided by both more standardization in and greater use of emerging, smarter approaches to evaluation.

Our collective knowledge will increase faster if the results from different sites and studies can be compared. Unfortunately, if results and health
outcome measures from different studies are not defined in the same way (standardized), their outcomes often cannot be compared—a significant problem unless you really like apples and oranges. For example, “prevented hospital emergency room visits” is a compelling outcome measure for housing and social services, but when there isn’t agreement on what health outcomes are preventable in the first place, it makes it difficult to determine the relative cost-effectiveness of different approaches. This lack of standardization is a significant problem in public health, community-based interventions, and the Community Guide, which we mentioned earlier, has struggled to identify the cost-effectiveness of different community-level preventive approaches because of the differences in how outcomes have been measured across studies.

Another benefit of standardization is that results can be aggregated across many evaluations when a standard definition and data collection protocols for outcomes and interventions are used. The federally funded Community Transformation Grant Program, which King County also participated in, required sites across the country to report activities and results using the same online data collection template. This uniformity provides evaluators with a much stronger ability to understand the effects of different approaches and allows findings to be analyzed much more robustly than if each of the more than 100 sites could only be compared against itself. New partnerships, such as the Build Healthy Places Network, offer promising efforts to standardize this new field as well.

Just as important as standardization is overcoming barriers to effective data collection. It is usually much less expensive to use information that someone else has already paid to collect. Examples include ongoing data collection efforts involving state health care claims data and public health disease-specific registries. An important obstacle, however, is the general lack of data on health at the neighborhood level. In general, community development work touches the lives of hundreds or perhaps thousands of people in specific neighborhoods. Existing health surveys often do not have an adequate sample sizes to offer meaningful or stable estimates in small areas. Instead, state or perhaps county or city estimates are typically as low as you can go. For
example, one of the best measurement systems, the Behavioral Risk Factor Surveillance System (BRFSS) survey, a telephone survey that is the basis of most communities’ knowledge of risk factors such as diet, physical activity, and tobacco use, was designed to report state-level results. The BRFSS requires additional local funding to produce local results for counties or cities, and additional funding still would be needed to measure outcomes from community development projects in a single neighborhood.

In King County, evaluators used standard methods and data across project activities. For example, data were used from BRFSS survey as well as the Healthy Youth Survey. In some cases, project funds were used to oversample in targeted geographic areas to ensure adequate sample size to detect a difference.

Careful thinking about the best measures to use, and in particular selecting outcomes that are common, can help avoid part of this problem. A neighborhood intervention to improve pregnancy outcomes might be hard pressed from a statistical standpoint to identify a decrease in infant mortality (with a base rate of a few deaths for every 10,000 births), but reductions in low birth weight (among every 100 births) might be feasible. Additional solutions to this problem of sample size limited by small geography may come from new information technologies. Geographic Information Systems (GIS) capability and ease of use are rapidly improving. Websites such as County Health Rankings and Community Commons⁶ allow users to trace the ZIP code and census tract distribution of social determinants of health and health outcomes in ways that were not possible in the past. New requirements on nonprofit hospitals to assess and report on the health profiles of the communities they serve has led to progress in the capacity to map and track changes in social determinants in smaller geographic areas through resources such as CHNA.org. In addition, analytic techniques such as data smoothing, which uses information from nearby census tracts, and the use of multiple-year rolling averages, can help build stable small area estimates for health and social determinants measures. In the future, as health outcomes become available through electronic health records, registries of specific health conditions and mapping

will become more useful in understanding the complex pathways from community conditions to disease and disability (while ensuring the privacy of individual health information).

Minneapolis/St Paul’s “Hennepin Health” is an example of a health system that has invested in social determinants because they have access to timely and actionable data. Hennepin Health is a county health plan for 10,000 of the highest need residents. The project spent considerable time and expense to create a near-to-real-time data warehouse for all the health services used by enrollees. When the Hennepin Health physicians could see a complete picture of all recent services, they observed multiple emergency room, physician, and pharmacy visits that did not contribute to optimal health nor efficient use of services. Using the data warehouse information, they were able to coordinate care and use the resulting shared savings to invest in social determinants like supportive housing and a sobering center.

There are also new tools to help think about and see data, such as network mapping and data visualization. Network maps that illustrate the connections between many organizations and the clients they serve can identify where duplication, fragmentation, and gaps in the system occur and how the entire system could benefit from specific policy changes. New data visualization tools allow policymakers to see at a glance complex relations in large data sets. For example, the Institute for Health Metrics and Evaluation’s Global Burden of Disease project (www.healthdata.org) has interactive online resources that use shape, size, and color to instantly and clearly show relationships of risk factors such as diet, sedentary lifestyle, tobacco use, and mental health to leading causes of death and disability over time and by location. As data visualization advances to show effects at a smaller scale, local policymakers will have a more data driven basis for decision making.

**MAXIMIZE INNOVATION THROUGH COLLABORATION**

Community development and public health are natural partners because both are focused on practical ways to improve the lives of people in their communities. Both are also interested in simple, inexpensive evaluation strategies. That said, the cross-disciplinary approach to evaluating how community development projects affect social determinants of
health is in its infancy. More tools are needed to aid collaboration. For example, a national clearinghouse that provides timely access to best practices and evaluation findings would allow those of us working on community development and public health to advance more quickly than separate efforts working in isolation on similar problems. We should be working collectively to develop logic models and establish causal connections between social determinants of health and health outcomes so these links can be explored and replicated, and evidence can be established in this new territory. We also should create mechanisms to standardize definitions and measures, particularly for community- as opposed to individual-level health determinants, such as availability of healthy food, open space, access to child-care, and others.

The King County project brought together experts across many sectors to develop and implement evidence-based projects to improve the community’s policies, systems, and the environment that affect health. Each sector has its own research and language; it took much work to just begin to be able to talk and share best practices with one another. Had we not had experts in so many fields, it would have been impossible to break into other disciplines to locate best practices. For example, few public health professionals are deeply engaged in the business of school siting, yet we know that schools within walkable distances of homes will encourage physical activity. A national clearinghouse of best practices would have been useful to streamline this work.

Knowing what works also takes investment, and evaluation needs resources. For example, the CDC recommends that 10 percent of tobacco prevention grant program budgets should be allocated to evaluation activities—this for a health issue with a relatively well-established evidence base.\(^7\) Community development is a $100 billion effort annually. The Department of Housing and Urban Development’s (HUD) budget of $45 billion includes 0.3 percent for research, evaluation, and demonstration projects.\(^8\) In short, the needed financing for measuring


the social determinants of health is not yet available, and we need a strategy to develop it.

The field also needs more cross-sector agreement on the concept of “return on investment” (ROI) as it relates to health measurements. In the strictest sense, of course, ROI is a monetary return for dollars invested. Some health interventions, particularly some prevention interventions (immunizations, asthma management, and tobacco cessation), do yield a financial return to the health care system. Another type of health intervention (for example, nurse home visiting for high-risk infants) has a monetary ROI, although the investor (public health) is different from the party reaping the economic benefit (criminal justice and economic sectors). Health improvements resulting from community development activities would most likely fall into this second category—improving multiple health outcomes, but with the return not necessarily to the original community development investors. Evaluations should be constructed to capture this second set of returns as well.

The public health world has taken this concept of ROI one step further, recognizing that health itself has a monetary value. As a consequence, often the ROI on health investments is not reported as dollars, but as health benefit for dollars invested. Increasingly, this metric is being standardized as a “healthy year lived” or “disability adjusted life year” (DALY), a measure that combines benefits from both reduced mortality and reduced morbidity. For example, childhood immunizations cost \$7 per added healthy year lived, while heart surgery costs \$37,000 per additional healthy year lived.\(^9\) Optimally, the cost per healthy year lived or DALY gained by investments in community development will be calculated and compared to costs from medical interventions. Being able to quantify health returns in this fashion may bring more investors to the table to make community improvements.

**SUMMARY**

Monitoring the effects of upstream determinants on health seems at first blush a straightforward task. But there are significant challenges that

---

must be overcome. We have suggested five approaches to lessen these challenges, as summarized below.

**Incorporate evaluation into project planning early on.** Incorporating evaluation early on increases the likelihood of a useful evaluation, may allow for more deliberate incorporation of health interventions, and could create opportunities for more elegant evaluation through smart project implementation.

**Clearly define causes and hoped for effects.** Then pick a practical and affordable evaluation design for measuring them. Clearly defining the intervention, hoped for effects, and if possible, a comparison group are key first steps. Randomized control trials are not the only option. Community developers have a co-conspirator in public health scientists who routinely conduct and evaluate similar kinds of interventions.

**Prioritize measuring what comes first rather than focusing on long-term outcomes.** Logic models are excellent tools for identifying the paths from inputs and activities to short-, medium-, and long-term outcomes. A logic model will help focus evaluation on measuring the first expected results so that program managers can know early whether they are on the right track. These “first results” with evidence-based links to longer-term health outcomes are among the most relevant to community development projects.

**Standardize, simplify, and innovate measurement methods and approaches.** Using simple and standard measurement methods will accelerate progress for both the community development and health fields and will allow for more accurate comparison across projects. Using existing health data is one strategy, but it is hampered by limited information at the neighborhood level. Selecting common outcomes, more effective use of GIS and analytic methods that identify health outcomes at smaller geographic areas, and the potential use of electronic health care records are all positive steps.

**Maximize innovation through collaboration.** The field is early in the process of figuring out how to effectively measure the social determinants of health, and inventing wheels is easier when you collaborate. A national clearinghouse of evidence would speed good intervention
design and evaluations and assemble known links in logic models. Given
the size of the investment this country is making in community develop-
ment, identifying resources for sound evaluation must be a priority.
Work is also needed to develop shared understanding of the concept
of return on investment, including reporting on this return in not only
dollars but also in health gains.

We are at the beginning of an interdisciplinary collaboration that has
the potential to increase the effectiveness of both the community devel-
opment and the health fields. Working together in smart ways will move
us forward quickly, and signs of early successes could create momentum
for greater investment in neighborhood-level improvements from new
sources, such as health care payers, insurers, and hospital community
benefit programs. As evidence (and our sophistication in obtaining
it) grows, it should enable the goal of using measures of community
features (such as grocery stores, bike paths and health clinics) to
accurately predict and improve both health and economic outcomes.

- - -

DAVID FLEMING, MD is health officer of Public Health—Seattle & King County, a large
Metropolitan Health Department that serves 2 million residents. Dr. Fleming has
also served as the director of global health strategies for the Bill and Melinda Gates
Foundation, the deputy director of the Centers for Disease Control and Prevention,
and the state epidemiologist for the Oregon Health Division. He has participated in
multiple national and international boards and advisory committees, has published on
a wide range of public health topics, and has faculty appointments at the University of
Washington and the Oregon Health and Science University.

- - -

HILARY KARASZ, PhD has been with Public Health—Seattle & King County since 2001.
She is currently a public information officer, where she works with the media, staff,
community members and others to provide high quality health information to King
County residents. She is also the director of a six-year CDC funded research project
that examines the potential for mobile health applications such as text messaging
to be used at the local public health department level. Hilary received a BA in History
from the University of California and a doctorate in Communication from the University
of Washington.
Kirsten Wysen, MHSA, has worked at Public Health—Seattle & King County since 2003. She is the initiative director of communities of opportunity, an early strategy of the King County Health & Human Services Transformation Plan. Before that, she analyzed Medicaid programs at the National Academy for State Health Policy in Portland, Maine. In the 1990s, Kirsten served as the Deputy Program Manager for the Washington Basic Health Plan and as a Policy Analyst at the state Health Services Commission and the Medicare Payment Assessment Commission in Washington, DC. She graduated from Brown University and the University of Michigan School of Public Health.