



HHS Public Access

Author manuscript

Health Aff (Millwood). Author manuscript; available in PMC 2018 March 12.

Published in final edited form as:

Health Aff (Millwood). 2018 January ; 37(1): 22–29. doi:10.1377/hlthaff.2017.1125.

Evaluating Community-Based Health Improvement Programs

Carrie E. Fry,

Doctoral student in health policy at the Harvard Graduate School of Arts and Sciences, in Cambridge, Massachusetts

Sayeh S. Nikpay,

Assistant professor in the Department of Health Policy at Vanderbilt University School of Medicine, in Nashville, Tennessee

Erika Leslie, and

Postdoctoral fellow in the Department of Health Policy at Vanderbilt University School of Medicine

Melinda B. Buntin

Professor in and chair of the Department of Health Policy at Vanderbilt University School of Medicine

Abstract

Increasingly, public and private resources are being dedicated to community-based health improvement programs. But evaluations of these programs typically rely on data about process and a pre-post study design without a comparison community. To better determine the association between the implementation of community-based health improvement programs and county-level health outcomes, we used publicly available data for the period 2002–06 to create a propensity-weighted set of controls for conducting multiple regression analyses. We found that the implementation of community-based health improvement programs was associated with a decrease of less than 0.15 percent in the rate of obesity, an even smaller decrease in the proportion of people reporting being in poor or fair health, and a smaller increase in the rate of smoking. None of these changes was significant. Additionally, program counties tended to have younger residents and higher rates of poverty and unemployment than nonprogram counties. These differences could be driving forces behind program implementation. To better evaluate health improvement programs, funders should provide guidance and expertise in measurement, data collection, and analytic strategies at the beginning of program implementation.

Over the past decade the private and public sectors have made large community-based investments in improving population health. Many of these investments have been made in multisector coalitions that seek to improve specific communitywide health outcomes, such as reductions in obesity or smoking. Through their programs, these coalitions develop consensus on targeted health outcomes, potential metrics, and programs for implementation; align existing resources in community-based organizations; and implement evidence-based interventions to fill programmatic gaps. Despite often substantial financial investment, little

is known about the relationship between the implementation of a health improvement program and the subsequent health status of the community.

Previous studies of community-based health improvement programs have found that they are influential in changing individual behavior and health-related community policies^{1,2} but do not produce significant changes in health outcomes, even after ten years.^{3–8} Much of the earlier literature that demonstrated positive changes in attributable health outcomes was limited to smaller, health care-oriented interventions, specific racial or ethnic groups, or highly specific health conditions.^{9–12} A more recent study investigating self-reported public health coalition activity found that greater planning activity was associated with reductions in mortality.¹³

These previous reports highlight the challenges inherent in evaluating community-based health improvement programs. Communities that implement these programs might not have sufficient resources to collect data or measure health outcomes. Evaluations of these programs typically rely on easily collectible data and pre-post designs without comparison or control communities. And because these evaluations do not adjust for secular trends, it is difficult to link program implementation to changes in health behavior, attitudes, or outcomes. Nevertheless, the economic and human capital investments being made in health improvement programs warrant the use of more rigorous research designs.¹⁴

This study used a pre-post design with county-level health status comparisons to evaluate community-based health improvement programs implemented in the period 2007–12. By combining multiple programs into a single analysis, examining changes in specific health outcomes, and using a more rigorous design, this study provides insight into such programs' potential to make positive changes in population health outcomes. Our analysis also demonstrates important threats to the validity of commonly used evaluation designs.

Study Data And Methods

Because many of the communities in our data set implemented programs at the county level, we focused on the association between these programs and county-level health outcomes. We used multiple sources of publicly available data to create an inverse propensity-weighted set of controls for conducting multiple regression analyses.

DATA

We conducted extensive internet searches for relevant community-based health improvement programs and contacted leaders at national foundations and governmental agencies engaged in population health efforts to identify an initial set of programs to examine. Through snowball-sampled conversations with these leaders and, subsequently, with leaders of the programs, we attempted to define the universe of programs that met our program criteria. (For the programs included in our analysis, see online appendix exhibit 1.)¹⁵ We shared this list with major foundations and agencies operating in this area to ensure that we identified all relevant programs, and we iterated our identification strategy based on their feedback.

We then defined the geographical areas (or program sites) covered by each implemented program. Most program sites involved only a single county, or a large metropolitan area within a county, but others encompassed multicounty regions. The majority of the programs were implemented at the county level, and programs serving areas larger than a county could be disaggregated to the county level, which suggested that county-level analysis was most appropriate for this study.

We included communities that implemented a program in the period 2007–12 if their program included multiple sectors, such as private industry, health care organizations, and public health departments; were externally funded; or received guidance, oversight, or technical assistance from a national coordinating agency. These selection criteria intentionally omitted many programs implemented by county or city health departments using federal or state grant money. Identifying programs in a less restrictive way would have introduced greater variability in the kind, intensity, and duration of the programs, which would have decreased the precision of the estimated effects and would have made it difficult to generalize findings to programs with specific characteristics.

We identified four programs implemented at fifty-two sites that collectively encompassed 396 counties (appendix exhibit 1 lists organization names and overall characteristics of the four programs included in our study).¹⁵ We classified each site by its foci (sites within programs could have different foci, and sites could also have multiple foci). Sites were classified as focusing either on overall health and well-being (two) or on specific health outcomes—namely, child health (six), tobacco control (twenty-three), diabetes (eight), obesity (thirty-eight), or other health outcomes (nineteen). Additionally, we identified each program's year of implementation, as well as the year of its termination (if applicable).

The outcome variables were county-level health outcomes obtained from the Selected Metropolitan/Micropolitan Area Risk Trends (SMART) data for the period 2002–12 from the Behavioral Risk Factor Surveillance System (BRFSS). BRFSS county-level SMART estimates are derived from metropolitan and micropolitan statistical areas (MMSAs) that have at least 500 respondents in a given year and 19 sample members in each MMSA-level stratification category (such as race, sex, or age groups).¹⁶ County-level estimates are weighted by procedures that employ known population demographics produced by the decennial census and American Community Survey.¹⁶ Over our study period, an average of 7.36 percent of US counties were included in the SMART data. The units of analysis for our study are county-year dyads.

We linked the SMART data and program data to county-level estimates of poverty and demographic and employment characteristics. Poverty data, including median household income and percentage living in poverty, were obtained from the Small Area Income and Poverty Estimates, produced annually by the Census Bureau.¹⁷ County-level age composition was obtained from Surveillance, Epidemiology, and End Results Program data, produced annually by the National Cancer Institute.¹⁸ Employment data were obtained from the Local Area Unemployment Statistics program of the Bureau of Labor Statistics.¹⁹

ANALYSES

Descriptive statistics of the number and type of community-based programs over time were produced. We then used inverse propensity score treatment weighting to reweight treatment and control counties. Regression analyses were conducted using a difference-in-differences design and an event study.

Our goal was to evaluate the implementation of any program, a tobacco-focused program, and an obesity-focused program. We examined programs that focused on tobacco and obesity separately because of the direct link between the implementation of these programs and changes in specific health outcomes captured in the SMART data. Additionally, we chose to focus on tobacco and obesity programs because of their growth in numbers over the study period. This growth was attributable, in part, to funding provided by the American Recovery and Reinvestment Act of 2009, which required a focus on tobacco control or obesity.

For each type of program, we were interested in the association between implementation and three county-level self-reported health outcomes: whether respondents reported being in poor or fair health, smoking status, and obesity status. We chose overall health because of the potential of any program to improve this outcome, and we chose smoking and obesity status because of our emphasis on tobacco- and obesity-focused programs. Programs that focused on other health priorities, such as diabetes and hypertension, may also improve smoking and obesity status, making the latter two outcomes relevant to a broader set of programs.

INVERSE PROPENSITY SCORE TREATMENT WEIGHTING

We employed inverse propensity score treatment weighting, using changes in pre-implementation covariates to reweight untreated counties to achieve greater balance on observed covariates and create a more appropriate control group.²⁰ We assessed the balance of observed covariates using standardized differences.²¹ These inverse propensity weights were then used in all subsequent regression analyses. (For details on the methodology, see the appendix.)¹⁵

DIFFERENCE-IN-DIFFERENCES ANALYSIS

We used difference-in-differences regression analysis to evaluate the association between the implementation of a health improvement program and county-level health outcomes.²² Because some of the counties in our data set were included in both the treatment and control groups, depending on the year of implementation, we also employed a difference-in-differences design in which only counties that did not implement a program during the study period were included in the control set. All regression models included county and year fixed effects. We clustered standard errors at the county level to address auto-correlation.

EVENT STUDY

To examine possible pretreatment trends in the study counties, we also used an “event study” design, which compared annual average outcomes for treated counties in each year leading up to and after the county implemented a health improvement program.^{23,24} Each of these

models also included county and year fixed effects, the same covariates that were included in our difference-in-differences analysis, and clustered standard errors at the county level.

SENSITIVITY ANALYSES

Communities that implemented a population health improvement program may be intrinsically different from communities that did not. This endogeneity presents a problem in the regression analyses above. One way to mitigate the potential biases attributed to endogeneity is to parse out programs where selection is less of an issue. Our data set included counties that were selected for the Communities Putting Prevention to Work program,²⁵ which was funded by the Centers for Disease Control and Prevention under the American Recovery and Reinvestment Act. Funding for this program was competitive, and communities that received funding had to demonstrate in their applications that they were “shovel ready” (that is, had developed the necessary coalition, infrastructure, or capacity to begin implementing evidence-based programs as soon as funding was obtained).

Additionally, there may be countercyclical effects on health resulting from the Great Recession (2007–09).²⁶ To address this concern, we excluded counties that received a Communities Putting Prevention to Work grant. These programs were implemented as a direct result of the recession, and counties that received these grants may have been more susceptible than other counties were to the countercyclical effects of the economic downturn.

All analyses were conducted using Stata, version 15. The Vanderbilt University Institutional Review Board considered this study exempt from review, based on its use of publicly available data.

LIMITATIONS

This study, like many quasi-experimental studies, had several limitations. First, counties with a health improvement program have economic and demographic characteristics that differ significantly from those of counties without such a program. Because these differences could be related to both the health outcomes of interest and the probability of treatment, our estimates could be biased. However, when we limited our analyses to programs that had received funding through the American Recovery and Reinvestment Act, a group arguably less subject to selection bias than the group of programs that was not competitively selected, we found that programs funded through the act were not associated with significantly different changes in county-level health or smoking status or obesity when compared to programs that had not received funding through the act.

Second, although the list of programs, their foci, and their years of implementation have been validated by the program staff of funders in this area, including large nonprofit organizations and governmental agencies, there is still a possibility that some unpublicized programs were excluded from this analysis. Additionally, we did not measure the intensity (that is, the number of interventions implemented or the number of people reached) of the implemented programs or the amount of financial resources invested. Failure to capture variations in these programs could also mask the true effects of larger, more resourced, or better-administered programs.

Third, programs could have different effects depending on the baseline levels of health conditions or behaviors. For example, we found some evidence to suggest that among counties with higher baseline rates of people who reported poor or fair health, implementation of a health improvement program was associated with significant decreases in the proportion of residents reporting such health. This type of analysis was beyond the scope of this study, but it merits further investigation.

Fourth, while our identification and classification strategy included the stated health outcome foci of these programs, we did not necessarily capture the full range of intended outcomes. For some communities, the intended outcome of the health improvement program could be changes to policies or procedures; for others, the goal could have been improvements in health education and knowledge or changes in health behaviors and outcomes. While all of these policies and programs may eventually lead to changes in health outcomes, such changes might not be the only or best source of measurement for all programs. Despite the validation of our selection criteria and the use of small-area estimates for health outcomes, obtaining adequate data for the evaluation of programs was difficult.

Finally, small-area estimates from the BRFSS SMART data are known to have measurement error, which could result in inflated standard errors. Thus, relying on existing sources of aggregate data would be problematic even for communities that may conduct more rigorous evaluations of their programs in the future. Additional data gathering for evaluation from both implementation and non-implementation counties may be necessary and could prove to be a challenge, in terms of both the quality of the data and the time and resources required. Despite these limitations, this study used the best data and most rigorous methods available to estimate the relationship between health program implementation and county-level health outcomes.

Study Results

The number of health improvement program sites grew substantially over the study period, from fourteen in 2007 to fifty-two in 2012. The number of counties with a health improvement program also grew, from 319 in 2007 to 396 in 2012. Before 2010, most of the programmatic sites were focused on child health or other health priorities. With the start of funding through the American Recovery and Reinvestment Act in 2010, the number of tobacco- and obesity-focused sites grew substantially, from one each in 2007 to twenty-four and thirty-seven, respectively, in 2012 (exhibit 1). While the relative share of programs that focused on hypertension, child health, and other health priorities decreased after 2009, the absolute number of these programs either remained the same or grew.

Before the implementation of any health improvement program (that is, in 2002–06), there were significant differences between counties that did and did not implement a program in the period 2007–12. Compared to non-implementing counties, the counties with a health improvement program had a larger share of young adults (ages 20–39) but a smaller proportion of nonelderly adults (ages 40–64) (exhibit 2). Additionally, counties with a program had significantly higher proportions of their populations living in poverty and higher rates of unemployment. Our inverse propensity treatment reweighting, however,

achieved balance among observable covariates in the pre-implementation period (appendix exhibit 3).¹⁵

DIFFERENCE-IN-DIFFERENCES ANALYSIS

Using a standard difference-in-differences analysis, we found that the implementation of a health improvement program was associated with a mean reduction of less than 0.06 percentage points in the population that reported being in poor or fair health and a mean reduction of less than 0.15 percentage points in the population that is overweight or obese (exhibit 3). However, neither of these results was significant ($\alpha = 0.05$).

Reweighting control counties with inverse propensity treatment score weights resulted in a reduction of more than 0.06 percentage points in the proportion of a county's population that was overweight or obese (exhibit 3). However, this reweighting resulted in an increase of more than 0.1 percentage points in the proportion of the population reporting being in poor or fair health. (For full regression output of the difference-in-differences analysis, see appendix exhibits 3 and 4.)¹⁵ As was the case with the unweighted difference-in-differences approach, these changes were not significant. In both difference-in-differences analyses, the implementation of a health improvement program was associated with an increase (greater than 0.03 percentage point and 0.05 percentage point, respectively) in the proportion of people who smoked (exhibit 3). Results from our event study analysis were substantively similar to the results from the inverse propensity treatment score weighting analysis. (For results of the event study analysis, see appendix exhibits 5 and 6.)¹⁵

The implementation of a Communities Putting Prevention to Work program program funded through the American Recovery and Reinvestment Act was associated with an average decrease of 0.05 percentage points in the proportion of the population that reported being in poor or fair health, compared to counties that did not implement a program (exhibit 4). Implementation of a program funded through the act was also associated with a reduction of greater than 0.18 percentage points in county-level rates of obesity or overweight. Similar to the implementation of all programs, the implementation of a program funded by the act was associated with an increase of less than 0.2 percentage points in the proportion of the population that smoked, though these changes were not significant. Restricting our analysis to programs not funded through the act produced results similar to those seen in our analysis of all programs (exhibit 3).

When we restricted the treatment group to programs that focused specifically on tobacco or obesity, we found that the implementation of a tobacco-focused program was associated with a reduction of less than 0.5 percentage points in the population that reported being in poor or fair health (exhibit 4). Additionally, the implementation of an obesity-focused program was associated with a modest and nonsignificant decrease (of 0.22 percentage points) in the proportion of people who reported being in poor or fair health. The implementation of a tobacco program was associated with the largest percentage-point reduction (0.20) in the proportion of the population that smoked, and, similarly, the implementation of an obesity-focused program was associated with the largest reductions (0.40 percentage points) in the proportion of the population that was overweight or obese. However, tobacco-focused programs were associated with an increase of roughly 0.10

percentage points in the proportion of people who were overweight or obese, and the implementation of an obesity-focused program was associated with an increase of less than 0.25 percentage points in the proportion of the population that reported smoking. None of these changes was significant.

Discussion

Our work provides modest evidence for the role of health improvement programs in improving certain health outcomes and also provides insights into the kinds of communities that have engaged in community-based health improvement efforts.

Program implementation was associated with modest reductions in the percentage of the population that reported being in poor or fair health or being overweight or obese, although these differences were not significant. Programs that focused on a specific health outcome (for example, tobacco control and obesity) were associated with greater changes in these outcome, compared to all health improvement programs. However, it is important to note that programs that focused on obesity saw increases in tobacco use and programs that focused on tobacco control saw increases in obesity rates, which suggests that these programs may focus on one health outcome to the detriment of others.

In the pre-implementation study period, counties that implemented a health improvement program were more economically disadvantaged and had younger populations, compared to control counties. Taken together, these differences could be the impetus behind a community's decision to implement a health improvement program. If this is the case, such programs may improve overall health status, but not to a degree that overcomes other potential measures of social or economic disadvantage—such as educational attainment rates, the predominant industry, or median household income.

Until now, most of the evidence supporting multisectoral collaborations for health improvement comes from studies that used a simple pre-post design, comparing people who received the intervention's services before and after its intervention. This study, in contrast, used population health outcomes and employed regression techniques and inverse propensity treatment score weights to construct a control group. The advantage of a controlled design is that it lends support for the association between program implementation and population health outcomes.

For instance, the simple pre-post design used in many of the studies cited above would not have captured declining smoking rates nationally during our study period and could have inadvertently attributed changes in smoking status to program implementation. Additionally, the use of a controlled design allowed us to capture and account for other, non-health-related differences among the communities we examined.

Improving population-level health outcomes is difficult, and it takes time to “move the needle” on health outcomes. For example, a decrease of 0.5–1.0 percentage point in the rate of smoking per year may be the maximum change that a community could expect when implementing comprehensive tobacco control policies and programs. This means that in a community with an adult population of 500,000 and an adult smoking rate of 20 percent, a

program would need to change the smoking behavior of 500–1,000 adults in a single year to obtain a decrease of 0.5–1.0 percent. The level and intensity of programming required for this level of change might not be available to many communities, and almost a decade of programmatic implementation and evaluation might be required to produce changes of this magnitude. Thus, five years of post-implementation data (the maximum in our data set) might not provide enough time for changes in health outcomes to be realized, depending on the intensity and specificity of programming. Future research could extend our study period to more recent years, potentially providing the necessary lag time to observe changes in population-level health outcomes.

Conclusion

Retrospective evaluation of collaborative, multi-sector health improvement initiatives, including the health improvement programs evaluated here, is difficult. A preferable method of summative evaluation is for programs to be engaged in evaluation before, during, and after implementation. However, in many situations, organizations and coalitions that lead, develop, and implement a program have expertise in community outreach and organizing, implementation science, or evidence-based practices, rather than in program evaluation.

Thus, an evaluation team should be employed to provide guidance and expertise in measurement, data collection, and analytic strategies at the beginning of program implementation. Early entry of such a team allows for the identification of control communities, gathering of necessary pre-implementation data, and formative evaluations that lead to a summative evaluation.

However, resources are scarce, and many communities that engage in these efforts require private investment, grants, and public funds to implement their programs. There are often few resources remaining for an evaluation of any kind, much less an evaluation on the scale described here. Grant-making organizations and private-sector entities that invest in the implementation of programs could consider also providing resources to perform a thorough summative evaluation to adequately evaluate their return on investment. In addition, they may want to invest in more-robust data collection, not only for evaluation but also to target needs and guide implementation of population health improvement programs more broadly.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgments

This work was supported by the Robert Wood Johnson Foundation (Grant No. 77330). The authors thank Oktawia Wojcik and Caroline Young for their thoughtful comments on earlier versions of the article and support of this project more generally.

NOTES

1. Lv J, Liu QM, Ren YJ, He PP, Wang SF, Gao F, et al. A community-based multilevel intervention for smoking, physical activity, and diet: short-term findings from the Community Interventions for

- Health programme in Hangzhou, China. *J Epidemiol Community Health*. 2014; 68(4):333–9. [PubMed: 24297972]
2. Driscoll DL, Rupert DJ, Golin CE, McCormack LA, Sheridan SL, Welch BM, et al. Promoting prostate-specific antigen informed decision-making. Evaluating two community-level interventions. *Am J Prev Med*. 2008; 35(2):87–94. [PubMed: 18617077]
 3. Brand T, Pischke CR, Steenbock B, Schoenbach J, Poettgen S, Samkange-Zeeb F, et al. What works in community-based interventions promoting physical activity and healthy eating? A review of reviews. *Int J Environ Res Public Health*. 2014; 11(6):5866–88. [PubMed: 24886756]
 4. Kloek GC, van Lenthe FJ, van Nierop PW, Koelen MA, Mackenbach JP. Impact evaluation of a Dutch community intervention to improve health-related behaviour in deprived neighbourhoods. *Health Place*. 2006; 12(4):665–77. [PubMed: 16253541]
 5. Wolfenden L, Wyse R, Nichols M, Allender S, Millar L, McElduff P. A systematic review and meta-analysis of whole of community interventions to prevent excessive population weight gain. *Prev Med*. 2014; 62:193–200. [PubMed: 24518005]
 6. Kreuter MW, Lezin NA, Young LA. Evaluating community-based collaborative mechanisms: implications for practitioners. *Health Promot Pract*. 2000; 1(1):49–63.
 7. Merzel C, D’Afflitti J. Reconsidering community-based health promotion: promise, performance, and potential. *Am J Public Health*. 2003; 93(4):557–74. [PubMed: 12660197]
 8. Roussos ST, Fawcett SB. A review of collaborative partnerships as a strategy for improving community health. *Annu Rev Public Health*. 2000; 21:369–402. [PubMed: 10884958]
 9. Liao Y, Tucker P, Siegel P, Liburd L, Giles WH. Decreasing disparity in cholesterol screening in minority communities—findings from the racial and ethnic approaches to community health 2010. *J Epidemiol Community Health*. 2010; 64(4):292–9. [PubMed: 19666632]
 10. Cruz Y, Hernandez-Lane ME, Cohello JI, Bautista CT. The effectiveness of a community health program in improving diabetes knowledge in the Hispanic population: Salud y Bienestar (Health and Wellness). *J Community Health*. 2013; 38(6):1124–31. [PubMed: 23913103]
 11. Kent L, Morton D, Hurlow T, Rankin P, Hanna A, Diehl H. Long-term effectiveness of the community-based Complete Health Improvement Program (CHIP) lifestyle intervention: a cohort study. *BMJ Open*. 2013; 3(11):e003751.
 12. Landon BE, Hicks LS, O’Malley AJ, Lieu TA, Keegan T, McNeil BJ, et al. Improving the management of chronic disease at community health centers. *N Engl J Med*. 2007; 356(9):921–34. [PubMed: 17329699]
 13. Mays GP, Mamaril CB, Timsina LR. Preventable death rates fell where communities expanded population health activities through multisector networks. *Health Aff (Millwood)*. 2016; 35(11):2005–13. [PubMed: 27834240]
 14. Wolfenden L, Wiggers J. Strengthening the rigour of population-wide, community-based obesity prevention evaluations. *Public Health Nutr*. 2014; 17(2):407–21. [PubMed: 23157970]
 15. To access the appendix, click on the Details tab of the article online.
 16. Centers for Disease Control and Prevention. SMART: BRFSS frequently asked questions (FAQs) [Internet]. Atlanta (GA): CDC; [last updated 2013 Jul 12; cited 2017 Nov 27]. Available from: https://www.cdc.gov/brfss/smart/smart_faq.htm
 17. Census Bureau. Small Area Income and Poverty Estimates (SAIPE): data set. Washington (DC): Census Bureau;
 18. National Cancer Institute. Surveillance, Epidemiology, and End Results (SEER): data set. Bethesda (MD): National Cancer Institute;
 19. Bureau of Labor Statistics. Local Area Unemployment Statistics: database. Washington (DC): BLS;
 20. D’Agostino RB Jr. Propensity score methods for bias reduction in the comparison of a treatment to a non-randomized control group. *Stat Med*. 1998; 17(19):2265–81. [PubMed: 9802183]
 21. Austin PC, Stuart EA. Moving towards best practice when using inverse probability of treatment weighting (IPTW) using the propensity score to estimate causal treatment effects in observational studies. *Stat Med*. 2015; 34(28):3661–79. [PubMed: 26238958]
 22. Ashenfelter O, Card D. Using the longitudinal structure of earnings to estimate the effect of training programs. *Rev Econ Stat*. 1985; 67(4):648–60.

23. Bailey, MJ., Goodman-Bacon, A. The War on Poverty's experiment in public medicine: community health centers and the mortality of older Americans [Internet]. Cambridge (MA): National Bureau of Economic Research; 2014 Oct. [cited 2017 Nov 27]. (NBER Working Paper No. 20653 Available from: <http://www.nber.org/papers/w20653.pdf>)
24. Jacobson LS, LaLonde RJ, Sullivan DG. Earnings losses of displaced workers. *Am Econ Rev*. 1993; 83(4):685–709.
25. Centers for Disease Control and Prevention. Communities Putting Prevention to Work (2010-2012) [Internet]. Atlanta (GA): CDC; [last updated 2017 Mar 7; cited 2017 Nov 27]. Available from: <https://www.cdc.gov/nccdphp/dch/programs/communitiesputtingpreventiontowork/index.htm>
26. Ruhm CJ. Are recessions good for your health? *Q J Econ*. 2000; 115(2):617–50.

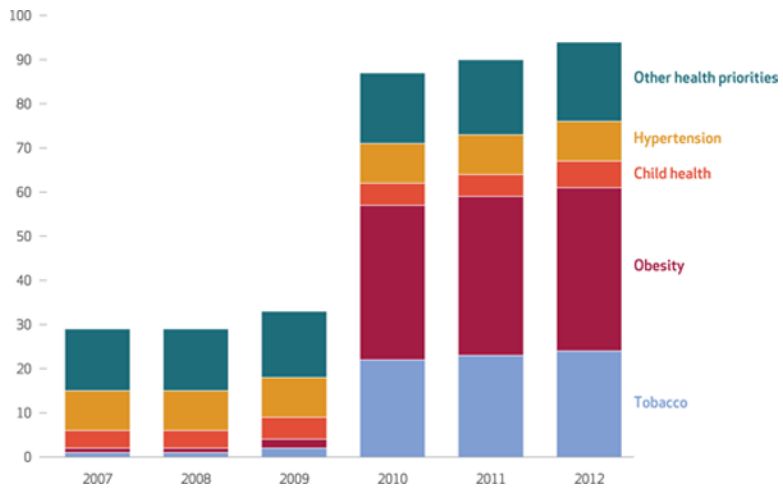


EXHIBIT 1. Numbers of community-based health improvement programs and their health outcome foci, 2007–12

SOURCE Authors’ analysis of selected community-based health improvement program data. **NOTES** The number of programs is cumulative over time. Programs with more than one focus are counted in each of their foci. Over this period, four programs were implemented at fifty-two sites that collectively contained 396 counties. The first program was implemented in 2007.

Author Manuscript

Author Manuscript

Author Manuscript

Author Manuscript

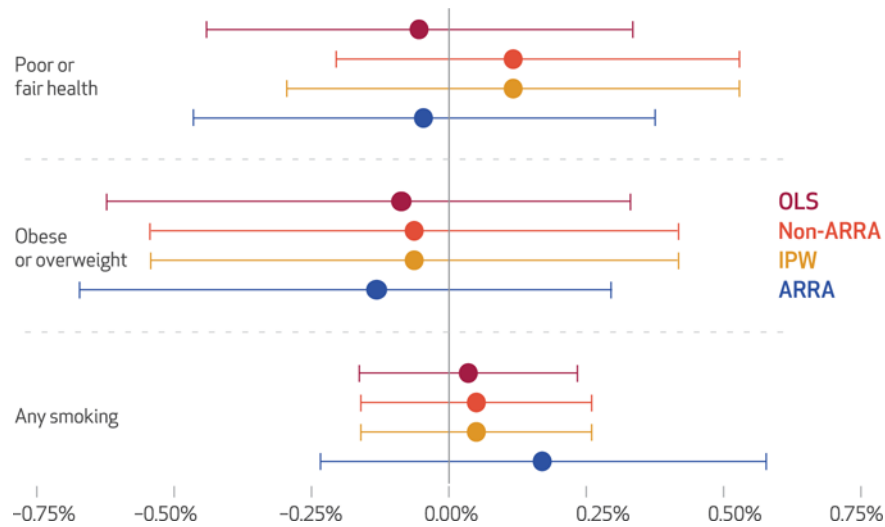


EXHIBIT 3. County-level changes in selected health outcomes after program implementation, by methodological approach

SOURCE Authors’ analysis of selected community-based health improvement program data and Behavioral Risk Factor Surveillance System Selected Metropolitan/Micropolitan Area Risk Trends data for 2002–12. **NOTES** The error bars indicate 95 percent confidence intervals. Models labeled “ARRA” include only counties that received funding via from the American Recovery and Reinvestment Act’s Communities Putting Prevention to Work grant. Models labeled “non-ARRA” include counties that did not receive funding from that grant. Standard difference-in-differences models are labeled “OLS” (ordinary least squares). Inverse propensity treatment score weighted models are labeled “IPW.” Statistical methods are described in the text, technical appendix, and appendix exhibits 2 and 3 (see note 15 in text).

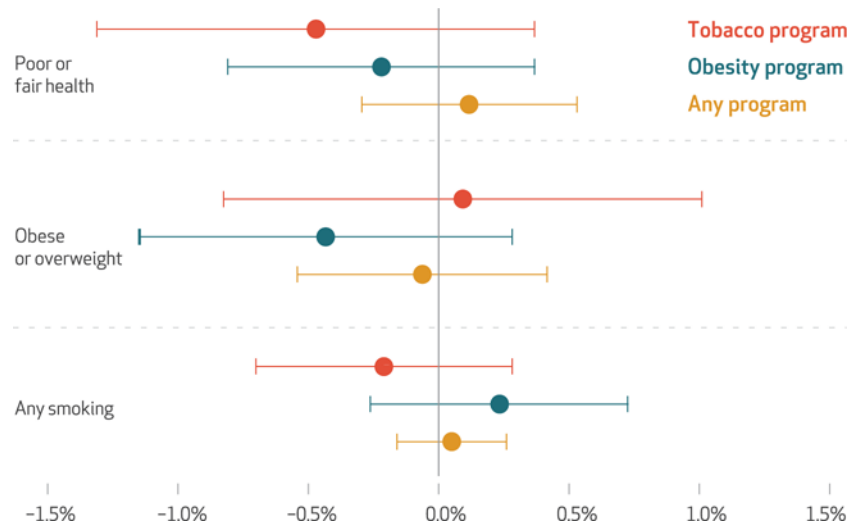


EXHIBIT 4. County-level changes in selected health outcomes after program implementation, by focus of program

SOURCE Authors’ analysis of selected community-based health improvement program data and Behavioral Risk Factor Surveillance System Selected Metropolitan/Micropolitan Area Risk Trends data for 2002–12. **NOTES** The error bars indicate 95 percent confidence intervals. “Any smoking” includes people who reported smoking daily and people who reported smoking some. “Obese or overweight” includes people with a body mass index of 25 to 40. Standard errors are clustered at the county (FIPS) level. Programs labeled as “obesity program” or “tobacco program” may also focus on additional health outcomes. All models are inverse propensity treatment weighted. Statistical methods are detailed in the text, technical appendix, and appendix exhibit 3 (see note 15 in text).

EXHIBIT 2

Sample characteristics of counties in 2002–06, before the implementation of health improvement programs, by implementation status

Characteristic	Non-implementing counties (<i>n</i> = 695)	Implementing counties (<i>n</i> = 269)
Population age range (years)		
0–19	27.91%	27.63%
20–39	27.50	28.66 ***
40–64	32.69	32.01 ***
65 and older	11.90	11.69
Population living in poverty	10.96%	13.34% ***
Unemployment rate	4.81	5.62 ***

SOURCE Authors' analysis of selected community health improvement program data and county-level Behavioral Risk Factor Surveillance System (BRFSS) Selected Metropolitan/Micropolitan Area Risk Trends (SMART) data for 2002–06. **NOTES** Counties with health improvement programs were not included in this analysis if there were fewer than 500 respondents in the BRFSS SMART data. Means were compared using unpooled *t*-tests of means.

p < 0:01