WIN-WIN PROJECT METHODOLOGY
EXECUTIVE SUMMARY

UCLA CENTER FOR HEALTH ADVANCEMENT

Contents

This document is intended to provide an overview of the methods utilized by the UCLA Center for Health Advancement to model outcomes for policies and programs. Additional technical details will be released in forthcoming academic papers, and details pertaining to individual interventions will be summarized in separate documents for each model.

The Win-Win project is a long-term initiative of the Center for Health Advancement at the Fielding School of Public Health, UCLA. It provides good science that drives real-world policy change by showing the health impact to populations and value to governments of policies, systems, and programmatic innovations. The project provides a standardized, unbiased economic analysis of interventions to help public-health officials make informed policy decisions and engage in cross-sectoral collaboration.

1. Preliminary Search

The first goal of the project was to catalog the full universe of program and policy interventions that utilize intersectoral means to achieve positive health impacts. We reviewed nine clearinghouses to collect names of and information on public health intervention programs and policies. Clearinghouses of evidence-based programs and policies list dozens of interventions for which there is strong evidence of effectiveness, and many of these are highly cost effective. These sources are described in **Table 1** in alphabetical order. Each source provides distinct methods for program and policy evaluation. At a minimum, we extracted program names and descriptions from each source. Additional information (available from some sources, but not all) included suggested outcomes, population served, estimated timeline for effectiveness, readiness for implementation, number of published studies, the quality of the research, the probability of a positive return, and costs to the government and/or participants.

The Washington State Institute for Public Policy (WSIPP) reports the following data for a given intervention: 1) detailed benefit estimates (taxpayer, participant, indirect, and other); 2) detailed cost estimates; 3) cumulative cash flows; 4) meta-analysis of program effects; and 5) odds of positive net present value. Based on these data, we created a ranking system using four criteria: 1) years to reach full effectiveness; 2) number of distinct studies that have estimated an effect for the intervention; 3) statistical significance and magnitude of effect sizes; and 4) the probability of a positive return on investment. Each of the criteria was populated for programs evaluated by WSIPP (not all programs listed by WSIPP were evaluated); however, other sources did not provide the same depth or breadth of information. The ranking system was completed to the extent possible in these instances.

The criteria were measured on a scale of one to four–with four being the most affirmative–based on information provided by the clearinghouse. The criteria were then averaged to arrive at an overall ranking for the program or policy. The timeline criteria was rated as a four if the program or policy was shown to have immediate cost savings, a three if it had short-term cost savings (up to four years), a two for intermediate cost savings (five to 15 years) and a one for long-term cost savings (16 or more years). The number of effect size criteria were rated based on the following: a four was given for programs or policies with 11 or more studies; a three was given for programs or policies with six to 10 studies; a two was given for programs or policies with three to five studies; and a one was given to programs or policies with one or two studies. For significance of effect sizes, a four was given for p-values less than .05; a three was given for p-values less than .10; a two was given for p-values less than .35; and a one was given for p-values greater than or equal to .35. Finally, the probability of positive returns was ranked according to the following: a four was given for probabilities between 80% and 100%; a three was given for probabilities between 60% and 80%; a two was given for probabilities between 10% and 60%; and a one was given for probabilities of less than 10%.

The cutoffs for the varying levels of ranking in each category were based on an analysis of the natural break points between clusters of studies and on more standard definitions of statistical significance.

When a database such as WSIPP provided meta-analyses on more than one outcome, the outcome and associated p-value with the most effect sizes was used as input for the ranking. Each program was then ranked among the other programs evaluated by the same clearinghouse. In this regard, a program with a thorough analysis from WSIPP was not ranked against a program analyzed by a less methodical clearinghouse. In addition, certain clearinghouses provided distinct information, and this unique data was used in the most appropriate manner in the ranking system. For instance, SAMHSA's National Registry of Evidence-based Programs and Practices (NREPP) provided estimates of readiness for implementation. This information was used instead of the timeline information from WSIPP in ranking the programs and policies from the NREPP. Other than WSIPP, SAMHSA, County Health Rankings and Roadmaps, and The Guide to Community Preventive Services supplied information that could be used to rank programs and policies. The other sources provided limited evaluation information, but the programs/policies were listed, described, and evaluated to the degree reasonably possible. The database of programs and the associated evaluation criteria will be updated periodically.

2. Methods for Estimating Effect Sizes of Interventions

Utilization of effect sizes varies by intervention. Specific calculations will be explained in the intervention-specific methodology reports, but in general, effect sizes are utilized in three ways. There is a hierarchy of effect sizes. We begin with the most preferred method and proceed to alternative methods as dictated by the source in the literature or by the available baseline data for the eligible population. Ranked from most preferred to least preferred effect sizes, they are:

(1) Weighted Standardized Effect Sizes based on Cohen's d or the D-Cox transformation
(2) Relative Risk Reductions utilized as a percentage decrease
(3) Absolute Risk Reduction utilized as a change in outcome units

The research team performs literature reviews for the intervention of interest to obtain the effect sizes associated with outcomes of interest. We use a modified approach similar to that developed by the Community Guide. Search terms are used to obtain as many relevant literature resources as possible, and then the studies are analyzed to determine whether they 1) evaluate a similar intervention, 2) use a strong study design, and 3) measure outcomes of interest to the model. This process begins by analyzing the literature used by the clearinghouse that was responsible for our preliminary literature review finding the intervention. From those studies, we obtain a clearer picture of the landscape of the literature and use the meta-analytic methods to expand the literature base from which we draw to obtain the relative risk reductions. These studies also allow us to develop a conceptual model for each intervention.

Once the studies are cataloged, we extract the relevant information provided based on our conceptual model leading to outcomes of interest. We take the absolute difference in means, the group means for different study conditions, variance, standard deviation, 95% confidence interval, etc. This information is organized into a spreadsheet by the type of outcome monitored so that it can be combined to obtain one effect size for each portion of the model. The effects are combined across studies with the same outcome by inverse variance weighting the effects. The variance is either reported or calculated from the 95% confidence interval.

In the event that one outcome is measured across multiple time periods, we plot each measurement against time to monitor whether the effect increases or decreases with more time. This plays an important role in the distribution of effects over years. When the model runs and we distribute the cost offsets across future years, we use this information to distribute the savings.

When there are intermediate outcomes, such as elementary test scores, we can use additional effect sizes from the literature to connect those test scores to our outcomes of interest–in this case, high school graduation. However, in cases where that outcome is later measured directly, we opt for the directly measured outcome rather than attempting to model it out.

3. Utilization of Effect Sizes

It is important to ascertain the baseline conditions for the population to be treated by the intervention. From that baseline, we apply the effects to determine the overall impacts of that program or policy. As described in the estimation of effect sizes, the form of the baseline data determines how the effect sizes are operationalized.

When the data provides information on dichotomous variables, proportional rates, or a continuous variable for which the distribution and standard deviation of the measure is available, we utilized the standardized effect size from Cohen's d or the D-Cox estimation. This application yields a relative risk reduction, but it is dependent on the baseline rates for the eligible population within the jurisdiction of interest. This means that if we have baseline rates, say for crime, by zip code, then the relative risk reduction will similarly vary by zip code. This is an advantage over utilizing one average relative risk reduction across the entire eligible population. This methodology also allows us to compare similar, but not identical outcomes to establish one effect size.

When the outcome of interest is continuous, but there is no information on the standard deviation, we utilize a relative risk reduction. This allows us to apply an effect to the baseline rate for the smallest jurisdiction for which we can obtain data. For example, consider a literature review that indicates an intervention has a 20% relative reduction in the number of cigarettes smoked in the last 30 days. If the baseline data available indicates that the mean number of cigarettes smoked by the eligible population is X, but there is no information on the standard deviation around that mean, then we would say that the avoided cigarettes would be 0.2X. The standardized effect sizes require either a dichotomous variable or a standard deviation, so this relative risk reduction can be used when the baseline data do not provide those numbers.

The final option for utilizing effect sizes from the literature is to use absolute reductions. These are measured in the same units as the outcome measures. For example, in one model analyzing remediation of asthma triggers for children, no reliable baseline data exists for the number of school days that an asthmatic child misses due to their condition. In this case, we apply the absolute reduction calculated from the literature of N days. This procedure clearly requires an assumption

that the average reduction is constant across jurisdictions and across the baseline number of school days missed.

3.1. **Standardized Effect Sizes.** When utilizing effect sizes, the equation to calculate monetary returns from the effects on crime, health, and education is:

Continuous Measures

$$\begin{aligned}\text{cost offset} &= \text{(standardized effect size) * (standard deviation of the baseline rate) *} \\ &\quad \text{(take-up rate) * (eligible population) * (cost per unit)}\end{aligned}$$

$$\begin{aligned}\text{averted units} &= \text{(standardized effect size) * (standard deviation of the baseline rate) *} \\ &\quad \text{(take-up rate) * (eligible population)}\end{aligned}$$

Dichotomous Measures

$$\begin{aligned}\text{cost offset} &= \text{(take-up rate) * (eligible population) * (cost per unit) *} \\ &\quad \left[\frac{e^{(ES*1.65)} * Baseline}{1 - Baseline + Baseline * e^{(ES*1.65)}} - Baseline\right]\end{aligned}$$

3.2. **Relative Risk Reduction.** The general equation used to model effects with relative risk reductions on baseline rates of crime, health, and education is as follows:

$$\begin{aligned}\text{cost offset} &= \text{(risk reduction) * (baseline risk for treated population) *} \\ &\quad \text{(take-up rate) * (eligible population) * (cost per unit)}\end{aligned}$$

Where the risk reduction is determined as previously described based on the literature, the baseline data is at the desired jurisdictional level if available, the take-up rate indicates the percent of the eligible population who are assumed to participate in the intervention if eligible, the eligible population is the number of individuals at the jurisdictional level who meet the criteria as specified by the

intervention, and the cost per unit is the cost per person or household of the outcome of interest at the baseline time.

Similarly, the general equation for the reduction in absolute number of units is as follows:

$$\begin{aligned}\text{averted units} =&(\text{take-up rate}) *(\text{eligible population})\\ &* (\text{baseline rate})* (\text{risk reduction})\end{aligned}$$

4. Baseline Data

Baseline data are collected from various sources at the geographic level of interest. For most models, the target level is the zip-code. When that is not available, we look to larger jurisdictions to obtain the data: neighborhood, city, county, state, region, nation. In some cases there is a particular jurisdiction that is of interest for an intervention i.e. school districts for an educational program. Many of the demographic data come from the US Census bureau, and the more localized data come from local health, education or criminal justice departments.

5. Results

The outcomes of the Win-Win model are categorized into the four areas of focus: crime, education, health, and government savings. While most outcomes are ultimately monetized as cost offsets to the government, they are also reported in their original units. The model reports outcomes such as increases in high school graduation rates, reductions in grade repetition or special education, decreases in various chronic health conditions, and decreases in the crime rates. These results are then published in a one page summary page and in a full report of all outcome measures and monetary returns and shared on our website, http://winwin.uclacha.org/. The analysis demonstrates the concrete impact of an intervention from year one of implementation onward and provides a standardized way to compare interventions and measure their impact in a variety of geographical areas. The results can promote cross-sectoral collaboration and inform program and policy decision-making.

FIGURE 1. Program Evaluation Sources

Name URL	Description
Blueprints for Healthy Youth Development http://www.blueprintsprograms.com/	Blueprints is funded by the Annie E. Casey Foundation. With this funding, outcomes have been expanded to include not only problem behavior, but also education, emotional well-being, physical health, and positive relationships. Blueprints has also been rebranded as Blueprints for Healthy Youth Development. More than 1,250 programs have been reviewed, but only a small portion of them have been designated as model or promising programs based on their ability to effectively improve developmental outcomes in the areas of behavior, education, emotional well-being, health and positive relationships.
County Health Rankings & Roadmaps http://www.countyhealthrankings.org	The County Health Rankings & Roadmaps program is a collaboration between the Robert Wood Johnson Foundation and the University of Wisconsin Population Health Institute. We believe America can become a nation where getting healthy, staying healthy, and making sure our children grow up healthy are top priorities. We have a vision of an America where we all strive together to build a national culture of health that enables all in our diverse society to lead healthy lives, now and for generations to come.
Edutopia http://www.edutopia.org/	Edutopia is dedicated to transforming the learning process by helping educators implement the strategies below. These strategies -- and the educators who implement them -- are empowering students to think critically, access and analyze information, creatively problem solve, work collaboratively, and communicate with clarity and impact. Discover the resources, research, experts, and fellow Edutopia members who are changing our schools.
SAMHSA's National Registry of Evidence-based Programs and Practices (NREPP) http://nrepp.samhsa.gov/	The National Registry of Evidence-based Programs and Practices (NREPP) is a searchable online database of mental health and substance abuse interventions. All interventions in the registry have met NREPP's minimum requirements for review and have been independently assessed and rated for Quality of Research and Readiness for Dissemination.
The California Evidence-Based Clearinghouse for Child Welfare http://www.cebc4cw.org	The CEBC helps to identify and disseminate information regarding evidence-based practices relevant to child welfare. Evidence-based practices are those that have empirical research supporting their efficacy. The CEBC provides guidance on evidence-based practices to statewide agencies, counties, public and private organizations, and individuals. This guidance is provided in simple straightforward formats reducing the user's need to conduct literature searches, review extensive literature, or understand and critique research methodology.
The Center for Health and Health Care in Schools http://www.healthinschools.org/	CHHCS at George Washington University specializes in researching effective school-connected programs, policies and systems, using the school location as a place-based solution for improving children's overall health and school success.
The Guide to Community Preventive Services http://www.thecommunityguide.org	The community guide covers a wide variety of programs and policies. However, many of the topics are only analyzed at the very general level. For instance, they will analyze whether it is recommended to encourage healthy eating and exercise. However, they do not analyze specific programs or policies for implementing that. Because of this, many of the small analyses are not as useful for our purposes. Further, the presented analyses show more big-picture outcomes like "5% reduction in alcohol usage" rather than monetizing them.
University of Wisconsin Population Health Institute http://uwphi.pophealth.wisc.edu/	The Institute serves as a focal point for applied public health and health policy within the University of Wisconsin-Madison School of Medicine and Public Health as well as a bridge to public health and health policy practitioners in the state. We strive to provide objective and collaborative program evaluation services to assess and enhance effectiveness and promote fidelity of implementation of programs and policies.
Washington State Institute for Public Policy http://www.wsipp.wa.gov	Since the 1990s, the Washington State legislature has directed WSIPP to identify "evidence-based" policies. The goal is to provide Washington policymakers and budget writers with a list of well- researched public policies that can, with a high degree of certainty, lead to better statewide outcomes coupled with a more efficient use of taxpayer dollars.

FIGURE 2. Example Screenshot of Ranking Information

Source	Name of Program	Category	Consol Outcomes DropDown	Outcomes Assessed (# of Effect Sizes)	(Net Benefit) /Cost	Timeline	Program or Policy
WSIPP	State and district early childhood education programs	Pre-K to 12 Education	Scholastic performance, Behavioral disorder, Crime, Grad Rate, Grade Rep, Special Ed	Test Scores(17), grade repitition(4), Special edu.(3), Grad rate(2), Crime(1)	4.51	Pre-k, positive at yr 20	Program

Population served	Timeline	# of ES	Sig. of ES (P-value)	Odds (+ NPV)	P(pos. returns)	Average Overall Rating
Universal for age group or taregt low-income	1	4	4	0.83	4	3.25

Rank	How Promising	Total Costs	Gov't Returns	Gov't Ben/Cost	Participant Returns	Participant Ben/Cost
24	Very likely to have positive returns. Crowdout may be concern if many already have the service.	$(7,020.00)	$ 9,734	-1.39	$ 14,433	-2.06