eLearning Series

Evidence-based Malaria BCC: From Theory to Program Evaluation

Module 5: Evaluating Social and Behavior Change Communication – Handout

Module 5 of 5

Marc Boulay, PhD

Deputy Director of Research and Evaluation
Center for Communication Programs
Johns Hopkins Bloomberg School of Public Health

Learning objectives

At the end of this presentation, participants should:

- Understand and list at least two main questions that SBCC outcome evaluations strive to answer in order to help program planners learn.
- Be familiar with and describe at least three standard approaches used for evaluating SBCC interventions.
- Be aware of and describe two current state-of-the-art approaches for determining the causal attribution of SBCC.

Table of contents

Part 1: Purpose of evaluating SBCC

Hello, my name is Marc Boulay and I'm the deputy director of Research and Evaluation at the Center for Communication Programs housed within the Bloomberg School of Public Health at the Johns Hopkins University. This lecture is the fifth and final module in this eLearning series and in it, I'll be talking about approaches for evaluating and social behavior change communication programs, or you may know them as SBCC or BCC programs. Hopefully you'll be able to use the information in this lecture as you consider how to determine if your communications activities were successful. Thanks for joining me.

Before I start the lecture let me just add that this eLearning series was supported with funding from the United States Agency for International Development through the NetWorks project.

This lecture has three main learning objectives. First, we hope that participants in this lecture will understand the main purposes of outcome evaluations for SBCC programs and come away from this lecture with an awareness of at least two questions that SBCC outcome evaluation should answer. One primary question for these evaluations is whether or not the program was effective. Did it or did it not result in a change in behavior? Equally important is to learn how a program worked. We recognize that SBCC program messages influence behaviors indirectly through knowledge, attitudes and beliefs that drive behavioral decisions. Understanding the specific attitudes through which messages affected behavior is important since this helps take the lessons from a successful program and apply them elsewhere.

The second objective of this lecture is to familiarize you with some of the standard approaches that are used to evaluate SBCC interventions and we hope that participants will know of three standard approaches for evaluating these interventions.

Third, we want to introduce you to some additional approaches that are becoming more popular for SBCC evaluation because of the strength of the assertions that they are able to support. Depending on the resources available within your organization and the strength of the evidence you require, more standard approaches may be sufficient for your needs. However, we want you to be aware of what some might call the "state of the art" for SBCC outcome evaluation and be able to list some of these approaches.

The lecture will consist of four parts and will then conclude with a brief summary of the key points followed by a short list of additional resources you may want to consult for more information. Part 1 briefly describes SBCC evaluation and the purpose of the SBCC evaluation. Part 2 presents some standard approaches for evaluating these interventions. In Part 3, we introduce the use of propensity score matching in SBCC outcome evaluation. As you will see, the use of propensity score helps us to determine whether or not a SBCC program was effective. Part 4 illustrates how mediation analysis helps you answer the question of how the messages influence behaviors. We will conclude with a summary of the lecture and provide you with some additional resources.

Let's begin with the first part. What is the purpose of evaluating social and behavior change communication programs? Why do we do it? What do we hope to accomplish?

This figure illustrates how evaluation contributes to all phases of a SBCC intervention and provides a summary overview of the range of data that support these evaluation activities. Formative evaluation occurs in the pre-intervention period and is designed to identify the characteristics of the audience, such as their attitudes and beliefs and their use of various communication channels that may be useful in designing the SBCC intervention. Monitoring evaluation occurs during the intervention period and is designed to assess the functioning of the program activities. Are activities happening? How are they being received by the audience? Program implementers use this information to make program refinements. Outcome evaluation occurs in the post-intervention period and is designed to identify the effects of the program.

One thing worth noting is that this entire evaluation system ideally occurs as a loop, where outcome evaluations contribute to the formative evaluation for subsequent phases of a project.

The remainder of this talk will focus on outcome evaluation and specifically on outcome evaluation using population-based survey data.

Like formative evaluation and program monitoring, the purpose of an outcome evaluation is to learn. What are the main questions that we are trying to learn about with an outcome evaluation?

Of course, one question is: did the program work? For SBCC, we usually look at behaviors as our ultimate measure of effectiveness. For instance, if your SBCC program intends to increase bed net use, you will want to learn whether or not your message successfully persuaded people to use a bed net. In addition to learning whether or not the program worked, a good outcome evaluation also tries to learn how it worked or if the behavior did not change, why the program did not have an effect on behaviors?

These questions are particularly important for SBCC evaluations because these interventions indirectly influence behaviors through knowledge, attitudes and beliefs. If we learn that a program effectively persuaded people to start using bed nets, but we do not know which of these indirect pathways led from the program to the behavior. It is difficult to apply the lessons of that intervention to other places. For example, did people begin using bed nets because through SBCC program they gained new knowledge about the threat of malaria and how to prevent it with bed nets? These pathways are important to evaluate.

Similarly, if a SBCC program did not result in behavior change, it is important to know why it didn’t work so that the next SBCC program may be more effective. For instance, a program may have successfully increased knowledge about bed nets but unless higher knowledge of bed nets prompted people to use them more often, the SBCC program would have little effect on behavior.

In this example, the outcome evaluation might discover that perceptions of social norms are much more strongly associated with bed net use than knowledge and that the program had no effect on changing these perceptions of social norms regarding bed nets. Learning this, program developers can then identify approaches or messages that would more effectively address perceptions of social norms.

Now that we have discussed what SBCC outcome evaluation is and what questions this type of evaluation can answer, we will move on to discuss standard approaches for SBCC outcome evaluation. We’ll do this in Part 2.

Part 2: Standard approaches for SBCC outcome evaluation

In this part, we'll describe some standard approaches for answering the question, "Did the SBCC program work?"

When we talk about outcome evaluations and public health, our benchmark for an effective design is the randomized controlled trial, also known as the RCT. RCT is a study that recruits a group of people and then randomly allocates these people into two groups. People in one group receive the intervention, and the people in the second group do not. After a period of time, researchers then examine whether the outcome of interest is higher in the group that receive the intervention compared to the group that did not. Any differences between these two groups are recognized to be due to the effects of the intervention.

There are several challenges associated with RCTs. Participants recruited into the study may be different in some way and not representative of the overall population. It may also be difficult to keep information about the intervention from the non-intervention group, resulting in what is called contamination of the control group. When an RCT is well implemented, it provides very strong evidence of the effectiveness of the intervention being studied.

Why does the RCT provide such strong evidence? To answer that, we need to know a little bit about a concept called confounding.

A confounder is a factor that is related to both exposure to the intervention and to the outcome of interest. For example, let us assume that we observed a positive relationship between a mother's use of child health services and her use of a bed net. This relationship might imply that by encouraging women to take their children for child health services, we may also influence their bed net use. More educated women may be more likely to use child health services, and more educated women may also be more likely to use a bed net. In this example, education is a possible confounding variable because it is related to both the exposure—use of a child health services—and the outcome of interest—bed net use. The importance of confounding is that it provides an alternate explanation to the observed relationship between the intervention and the outcome. The observed relationship between the use of child health services and bed net use may simply reflect the role of education and not the influence of child health service utilization on subsequent bed net use.

By randomly allocating individuals to a treatment group or a control group, RCTs essentially create two groups that are identical in all respects except one—receipt of the intervention. Since no other characteristics will be associated with the outcome except exposure to the intervention, RCTs effectively exclude the possibility of a confounder providing an alternative explanation for the difference between the treatment and control groups. Remember, a confounder is a factor that is associated with both the outcome of interest and the exposure to the intervention. Lacking any alternative explanation and observed difference in the outcome between the treatment and control groups can be strongly attributed to the effects of the intervention.

Despite the strength of the RCT as a methodology for outcome evaluations, there are some reasons why they may be inappropriate for use in evaluations of SBCC interventions.

First, it is difficult to randomly allocate people to groups that do or do not receive SBCC messages. This is particularly true for interventions that broadcast messages using mass media channels such as radio or television. With these channels, all individuals in the broadcast area are able to receive these messages outside of the control of program implementers.

SBCC programs that rely solely on local community-based channels for the intervention would have the ability to randomly allocate communities, although not individuals, to receive an intervention. However, the number of communities that would need to be included in an RCT to ensure that randomization results in equivalent exposed and controlled groups would need to be very large.

The diffusion of program messages is the second reason that RCTs may not be appropriate for SBCC evaluations. Program messages may diffuse when individuals exposed to the messages communicate this health information to their family and friends in their social network, or when individuals from outside the study area travel into the study area. From the SBCC program implementer's perspective, this diffusion is desired and even encouraged as some interventions promote interpersonal communication about the message topics. From the RCT perspective, the diffusion of program messages is problematic, since it may result in members of a control group becoming exposed to the intervention. When that happens, comparisons between exposed and unexposed groups no longer serve as a measure of a program's effect, since individuals in the control group may also have an influence by the intervention.

The third challenge associated with using RCTs for SBCC evaluation is that they are designed to assess only whether a program is effective or not. They do not lend themselves to the type of analysis necessary to explore the reasons why a program worked, or why it might not have worked.

Because of these challenges, SBCC evaluations typically use alternative approaches.

One common approach is to collect survey data prior to the intervention using a baseline survey, and then again following the intervention using an endline survey. This approach somewhat mirrors the RCT design in that it relies on a comparison of similar groups of people, where one group is exposed to the intervention and the second group is unexposed. In this approach, the baseline survey sample comprises the unexposed group and the endline survey comprises the exposed group.

Using this approach, a comparison of the change and the outcome between the baseline and endline surveys constitutes the effect of the intervention. However, the strength of this comparison depends on certain assumptions that I will discuss shortly.

As an example of this approach, this graph presents data from the 2005 and 2009 evaluation surveys conducted by the Health Communication Partnership in Zambia. The 2005 survey was conducted prior to the SBCC activities and serves as a baseline sample, while the 2009 survey was conducted at the end of the SBCC activities, and serves as an endline. The SBCC activities included messages designed to promote bed net use among pregnant women and young children, appropriate treatment for fevers in young children, as well as intermittent presumptive treatment (IPTp) for malaria among pregnant women.

This graph shows the change in four behaviors between the two surveys, three of which increased significantly between 2005 and 2009. The percent of children under 5 years of age that slept under an insecticide-treated bed net the night prior to the survey rose from 30 to 36 percent. The percent of children under 5 with a fever in the past two weeks that received appropriate treatment rose from 27 percent in 2005 to 57 percent in 2009. The percent of pregnant women that slept under an ITN the previous night increased from 26 to 35 percent. There was no change in IPTp use among women who had given birth in the past two years.

According to this approach, this slide indicates that the SBCC program messages had a significant effect on the use of ITNs among young children and pregnant women, and on treatment-seeking for childhood fever.

There are some limitations to this approach that suggest we should we be cautious when interpreting this comparison as the effect of the intervention.

One limitation is that this approach assumes that the SBCC intervention was the primary driver of any change in the outcome between the baseline and endline surveys. While this assumption may be plausible in some circumstances, it will not be plausible in many others. In the present example, it is likely that the large influx of bed nets brought into Zambia with Global Fund support (while the SBCC program was ongoing) contributed to the increase in net use, while the increased access to combination therapies for malaria contributed to the rise in treatment-seeking between the baseline and endline surveys. However, since the primary comparison in this approach does not explicitly include exposure to the SBCC activities, it is difficult to separate the effects of the SBCC activities from other influences on behavior. The bias introduced by these other influences will lead to an overestimate of the program's effects when relying on a comparison between the baseline and endline surveys.

The second limitation with this approach is that it assumes that everyone in the endline survey is exposed to the intervention. While this may have been realistic when a single national radio station was the one source of information, and reached the majority of the population. In today's media environment, few SBCC programs reach more than 50 percent of the intended audience and many reach fewer than that. This means that at least half of the intervention group in this design—defined us the endline sample—will consist of individuals unexposed to the intervention, and among whom one would not expect any effect of the SBCC messages. This implies that the baseline-endline comparisons will substantially underestimate the effect of the messages, unless the reach of these messages is nearly universal in the endline survey.

In short, using a baseline and endline comparison to evaluate the effectiveness of SBCC messages will often include two biases—one that will overestimate the program's effects and another that will underestimate them. Together, these two biases make it difficult to interpret the change over time as the effect of the program.

In addition, and just like the RCT, this approach also is geared mainly to learn whether or not a program worked and has limited ability to assess how it worked or why it did not work.

Time series approaches adapt the baseline-endline comparison to provide a second alternative to the RCT for SBCC program evaluation.

By time series, I'm referring to data that are collected at multiple points in time and often using relatively short intervals of time between data collection. This method improves on the baseline-endline approach because it can more explicitly link the changes that occur over time to the SBCC intervention, and other potential influences occurring during the intervention period. It does this by allowing for a more precise measure of when the change and the outcome started to occur, and examines this change in relation to the timing of events, including the start of the SBCC activities. Abrupt changes that occur soon after the start of SBCC activities will be more plausibly connected to those activities rather than to other events that occurred at other points in the time series.

There are some challenges which limit the usefulness of this approach. The need for multiple and frequent data points makes it difficult to collect the necessary data using household surveys. Often, time series are only feasible when the outcome behavior involves the use of a commodity whose sales or deliveries can be tracked easily. However, sales and delivery figures do not directly measure the behavior at the individual level and may provide inexact measures of behavioral outcomes.

This graph illustrates an example of a time series approach for evaluating a SBCC intervention—the Stop AIDS Love Life campaign in Ghana. The bars in the graph show the number of condoms in millions sold in Ghana in six-month intervals, although these data actually use the number of condoms delivered from the central warehouse as a proxy for condom sales, which is itself a proxy for condom use. The dotted red line indicates when the SBCC activities began.

As the graph shows, the number of condoms delivered from the central warehouse was relatively constant for the four years preceding the SBCC intervention. (As an aside, let me point out that with a time series approach, it is recommended to have data for a long period of time prior to the start of the SBCC to establish the baseline trend.) Once the intervention began, the amount of condoms delivered from the central warehouse spiked from approximately 4 and a half million condoms delivered in the six months prior to the campaign to nearly 8 million condoms delivered in the six months following the start of the campaign. It continued to increase in each successive six-month period.

These data may suggest that the campaign stimulated increased demand for condoms in Ghana. Other factors may have also been influencing condom use in Ghana, but it is less likely that these factors would have been introduced at exactly the same time as the SBCC campaign. While these results appear compelling, it remains unclear whether and to what extent these condom sales figures reflect actual condom use.

It is worth pointing out that this approach is more difficult to use when the outcome behaviors do not reflect the commodity, or when the distribution of that commodity is not controlled by a few sources willing to provide the data. In this example, USAID was the source of almost all condoms in Ghana, and could provide the data relatively easily.

A third alternative to the RCT is to use self-reported exposure to SBCC messages in household surveys to construct the groups of exposed and unexposed individuals. In this approach, a series of questions in a household survey are constructed to ask each respondent about their exposure to SBCC messages and to specific program elements such as logos and slogans. The responses to these survey questions are then used to categorize individuals as either exposed or unexposed to the program messages. This approach more precisely matches the RCT approach of defining the groups based on their exposure to an intervention. Although in this approach, membership in the exposed and unexposed groups are defined by the individuals themselves based on their recall of prior exposure to the messages, rather than by randomly allocating individuals to groups prior to the intervention as is done in an RCT.

In this approach, the effect of SBCC messages is determined by comparing the difference between the exposed and unexposed groups.

This graph illustrates data from the 2009 HCP Zambia endline survey to provide an example of the exposed-unexposed comparison approach to SBCC outcome evaluation. It divides survey respondents in two ways. At the top, it divides respondents into groups that were exposed or unexposed to malaria messages during community-level events. At the bottom, it divides respondents into three groups based on the level of exposure to malaria messages from mass media channels—no exposure, low exposure and high exposure. It is actually preferable when possible to use this type of multi-level grouping to capture a dose-response effect of exposure. We were able to separate out community and media exposure since the two messaging sources were largely independent from each other. The bars for each group report the percent of children under 5 that slept under an insecticide-treated bed net the previous night.

There was no significant difference in bed net use among children when comparing mothers exposed to community-level messages about malaria with mothers unexposed to community-level messages. However, significant differences were observed when comparing the three media exposure groups. Children in households with low levels of media exposure to malaria messages were more likely to have slept under a bed net than children in households with no media exposure to malaria messages. Children in households with high media exposure to bed nets were more likely to have slept under a bed net than children in low media exposure households. Thus, a clear dose-response is apparent—the greater the dose of media exposure to malaria messages, the more likely children are to have slept under a bed net.

Using this approach would lead one to conclude that the community-level messages were not effective at promoting bed net use among children, while the mass media messages about malaria were effective.

The approach that we just described has many similarities to an RCT which should support its use as an effective approach for SBCC evaluation. Just like an RCT, it categorizes individuals based on their receipt of the intervention, and just like an RCT, the difference in the two groups should reflect the effect of the SBCC program assuming that there are no confounders. As a reminder, confounders are factors that are associated with both the exposure to the SBCC and to the outcome behavior and introduce alternative explanations for the apparent effects of the SBCC intervention.

There is one crucial difference between this approach and the RCT. In the RCT, individuals are randomly allocated to receive the intervention. In this approach, individuals self-select to receive the SBCC messages. In other words, there may be some characteristics of individuals—their education, their motivations, something unknown—that make them more likely than other people to report being exposed to the SBCC messages. If these same characteristics are also associated with the outcome behavior, we can no longer assume that no confounders exist and therefore, there may be alternative explanations why a behavior is more common among individuals exposed to the SBCC messages.

We can account for some potential confounders by including them in the analysis. In the last slide, we controlled for age, education and urban/rural residence when calculating the percentages of children that slept under a bed net. Controlling for variables in our analysis is one way to account for potential confounders. The problem is that we do not know if there are some additional factors that we have not measured that are also confounding the relationship between exposure and behavior.

Now we've come to the end of Part 2 of this lecture. We've discussed some standard approaches for evaluating SBCC programs, and now we'll take a brief pause. When we come back, I'll introduce some more current state-of-the-art approaches for addressing some of the limitations with these standard approaches.

Part 3: Propensity Score Matching Approaches for Outcome Evaluations of SBCC

Hi, this is Marc Boulay and we'll start now with Part 3 of the lecture where we introduce an approach that uses "propensity score matching" to account for the potential confounders, both measured and unmeasured, in order to remove the possibility of alternative explanations and therefore to increase our confidence that the observed effects can be realistically attributed to the SBCC intervention.

This approach uses data collected in post-intervention household-based surveys and uses two steps to address potential confounding.

First, we calculate each individual's propensity to be exposed to the SBCC messages using their background characteristics and then match exposed individuals with unexposed individuals with the same propensity to be exposed.

Second, we then use a dual equation system to assess the likelihood that there are unmeasured confounders of the relationship between SBCC message exposure and the outcome of interest.

If we are able to determine that there does not appear to be any unmeasured confounders, our comparison between matched, exposed and unexposed individuals will be equivalent to an RCT. Remember that the hallmark of an RCT is the lack of confounders and consequently the lack of any alternative explanations for a difference between the exposed and the unexposed groups.

Now let me discuss the process that we use to calculate each individual's propensity score in more detail. What is a propensity score? It is simply an individual's likelihood of being exposed to the SBCC messages based on their background characteristics such as age, education, wealth status. The key point here is that any measured variable that could not be caused by exposure to the program's messages can be used to calculate the propensity score. For instance, a person's education level might have influenced the person's exposure to the program messages, however, their education level could not have been influenced by their exposure to the program messages. But a person's knowledge of bed nets might have been influenced by message exposure and therefore should not be used to calculate the propensity score.

How do we calculate the propensity score? We simply use a logistic regression model that uses self-reported message exposure as the outcome variable and all of the background variables as the independent variables. The predicted probability of message exposure derived from this model is the propensity score for each individual.

Then the individuals in the sample are matched to form groups of exposed and unexposed individuals with the same propensity to be exposed. This is done by matching each exposed individual to one unexposed individual with the same propensity score. In this way, for every exposed individual you have an unexposed individual with the same background characteristics.

By matching on their propensity score, we essentially create two groups of individuals, one group of exposed and a second group of unexposed, who are identical with respect to the variables that we use to create the propensity score. Since these variables are no longer associated with exposure, it removes the possibility that they can confound the relationship between exposure and the outcome behaviors.

Although the first step removed many of the possible measured confounders, we still need to determine if there are any unmeasured or unknown confounders that may explain any relationship between exposure and behavior. This step will assess the likelihood that unmeasured confounders exist.

We do this by estimating two probit regression models simultaneously. A probit regression model is similar to a logistic regression model, although it has normally distributed error terms. One model is identical to the model that was used to calculate the propensity scores and has exposure as the dependent variable and background variables as the independent variables. The second model has the outcome behavior as the dependent variable and exposure to SBCC messages as the independent variable.

It is important to remember that the error (or residual) terms in regression models contain all of the remaining variance of the outcome variable that is not explained by the independent variables in the model. In other words, the error term contains all of the other unmeasured factors that explain the outcome variable. If the error term from the first model predicting exposure is correlated with the error term from the second model predicting behavior, that implies the presence of some factor influencing both outcomes and potentially confounding the relationship between exposure and behavior. On the other hand, if the error terms are not correlated, this implies that there is no factor that explains both and therefore no unmeasured confounder of the relationship between exposure and behavior. In this case, the comparison between the matched groups is equivalent to an RCT.

To illustrate this approach with an example, we use data collected as part of the 2010 Zambia Malaria Indicator Survey. Our outcome variable for this analysis is the variable recording whether or not each woman in the sample reported sleeping under an insecticide-treated bed net the previous night. As this chart shows, 74 percent of women reported sleeping under an ITN the previous night and 26 percent reported that they did not sleep under an ITN the previous night. There are two questions of interest to us. Did the SBCC messages influence this behavior? And how did the SBCC messages influence this behavior? The following slides will seek to answer these questions and will use the propensity score matching approach.

First to calculate the propensity scores, we used seven background variables—age, has a child under the age of 6, number of years of formal education, their household's wealth quintile, their province of residence, whether they live in an urban or rural area, and whether their district received indoor residual spraying for malaria.

One important point to note is that while these variables may have influenced a person's exposure to malaria messages, the reverse cannot be true. It is not possible that exposure to malaria messages influenced any of these variables. Therefore, we can use these variables to calculate each individual's propensity score of being exposed to the SBCC messages.

The second thing to note is that these seven variables explained 35 percent of the total variance of the exposure variable. For propensity scores our rule of thumb is that background variables should explain at least 20 percent of the variance so we feel comfortable with the propensity scores that were calculated with these data.

Using the propensity scores, we then matched pairs of respondents—one exposed to the SBCC interventions and one unexposed—using the nearest neighbor approach with the caliper set to 0.005. This means that we take an exposed respondent and find the unexposed individual who has the closest propensity score to that respondent and then match the two respondents together. The caliper means that the respondents will only be matched if their propensity scores are within 0.005 of each other. Also note that we matched with replacement meaning that each unexposed individual could be matched to more than one exposed individual.

This slide compares the background characteristics between the exposed and unexposed groups prior to and following matching. It really isn't necessary to look at each comparison in detail here. Rather, focus on the red ovals that indicate the p-values for these comparisons. Prior to matching, these background variables varied significantly between the exposed and unexposed groups. But once we match them on their propensity scores, the exposed and unexposed groups are nearly identical in terms of these background characteristics.

Since these background variables do not differ between the exposed and unexposed groups, the matching has effectively removed them as possible confounders when we compare the behavioral outcome, bed net use, between the matched sets of exposed and unexposed individuals.

Once we have used the propensity scores to create the matched pairs, we need to identify whether any unmeasured confounders exist. As I mentioned previously, we do this by using a two-equation system in which one equation predicts exposure to the SBCC messages and the second equation predicts bed net used the previous night.

After estimating these equations we calculate rho. Rho is the measure of the correlation between the error or residual terms in the model predicting exposure with the error or the residual term in the model predicting bed net use. If the two residual terms are correlated, it implies that there exists some factor that is associated with both exposure to the SBCC messages and with bed net use and that may confound the relationship between the two. But if the residual terms are uncorrelated, we can be confident that no factor exists that influences both SBCC exposure and bed net use and thus, no unmeasured confounder exists.

In this example, our measure of rho was equal to 0.035 and test that this correlation was different from 0 had a p-value equal to 0.656. In other words, the two residuals were uncorrelated with each other, implying that there are no unmeasured confounders for our evaluation comparison of bed net use by exposure to the SBCC messages.

Now, we are finally ready to examine whether the exposure to the SBCC messages influenced bed net use. This slide compares the proportion of women sleeping under an ITN the previous night for those exposed and unexposed to the messages. The comparison on the left is the unadjusted comparison which corresponds to the comparison that we would have made had we not used propensity scores to match the respondents. The comparison on the right is the comparison using the matched pairs.

Without any adjustments, exposure to the messages appears to have resulted in a 15.6 percentage point increase in the use of bed nets while the matched comparison indicates that exposure results in a 12.5 percentage point increase in bed net use. The difference in this effect size reflects that some of the unadjusted effect can be attributed to the confounding effects of the background variables used to create the propensity score. But even accounting for the effects on net use attributed to these confounders, SBCC exposure still accounts for a significant and substantial effect on bed net use. Since we have already determined that there are no unmeasured confounders of this relationship, we can conclude that the SBCC messages did increase net use among women living in Zambia. Lacking other confounders it is difficult to identify any alternative explanation to account for the difference in bed net use between the matched exposed and unexposed pairs.

This ends Part 3 of the lecture where we introduced propensity score matching as a relatively new approach for SBCC outcome evaluation. When we come back we'll introduce Part 4 of the lecture where we discuss mediation analysis and approach for examining how SBCC program messages worked.

Part 4: Mediation Analysis to Identify How a SBCC Program Changed Behavior

Hi, this is Marc Boulay. Now we'll start Part 4 of the lecture where we introduce the idea of mediation analysis as an approach for assessing how program messages worked.

By now, it is well know that SBCC messages do not influence behavior directly. Rather, these messages influence a range of ideational factors that include both cognitive developments such as attitudes, perceived norms, and perceived risk, emotional elements such as confidence in one's ability to perform a behavior, and social elements like support and peer influence. These ideational factors influence behavior by affecting how individuals make decisions related to a behavior. In other words, these ideational factors mediate the relationship between exposure to SBCC messages and behavior.

When done strategically, SBCC interventions identify the specific mediators that are most likely to influence the behavior in the context of the study. We refer to this as the theory of the program. Program messages are then crafted to target these mediators. You may recall that previous modules of this eLearning series discuss the use of theory in designing SBCC messages that target some of these mediators.

The outcome evaluation should incorporate this theory by testing these mediating pathways. In other words, we look to see which ideational factors were influenced by the communication program, and then which mediating factors influenced the behavioral outcome. By identifying these pathways we are able to illustrate how the communication messages influence the behavior.

This has relevance for subsequent programs seeking to change the same behavior elsewhere. We often focus on which communication channels we should use—mass media, community-based or facility-based. But most health communication practitioners would recommend using all of these channels to communicate messages. While we know which channels should be used, there is often limited or no guidance on what the messages should say. By highlighting these mediating pathways, we can offer guidance to other programs on the content of their messages delivered through various communication channels.

To illustrate this approach, we will return to our analysis of the 2010 Zambia Malaria Indicator Survey. We note at the onset that a mediation analysis can be difficult to do with an existing data set that was not designed to measure the mediating pathways. Consequently, it does not contain measures for many of the hypothesized mediating factors.

Using the current data set, we examine the role of one potential mediator—do women recognize mosquitoes as the sole cause of malaria? While most survey respondents indicate mosquitoes as a cause of malaria, people often indicate that malaria has other causes as well. These beliefs may limit bed net use if people think that they can get malaria through these other causes, whether or not they sleep under a bed net. We examined whether this belief mediated the relationship between exposure to the SBCC messages and bed net use.

Our initial approach was to examine the two steps of this mediation directly—was SBCC exposure associated with the belief that mosquitoes are the sole cause of malaria? And then, was this belief associated with bed net use?

The graph on the left compares the percent of women that report that mosquitoes are the sole cause of malaria among those exposed and not exposed to the SBCC messages. This belief was common in both groups, but it was also positively associated with SBCC exposure. Women exposed to SBCC messages were significantly more likely than unexposed women to believe that mosquitoes are the sole cause of malaria.

The graph on the right examines the second step in this pathway—is belief that mosquitoes are the sole cause of malaria associated with bed net use? Women who believed that mosquitoes are the sole cause of malaria were significantly more likely than women who believed that there are other causes to report sleeping under a bed net the previous night.

Overall, this slide indicates that both steps in the pathway are significant and suggests that this belief may mediate the relationship between exposure and behavior.

To look at this mediation analysis more systematically we adopted the approach recommended by Baron and Kenny. This approach first runs a regression model using the exposure variable to predict bed net use to calculate the total effect of the SBCC messages on bed net use. Then, a second regression model is used that includes both the exposure variable and the potential mediator variable to predict bed net use. If the exposure variable is mediated through our belief variable, the size of the association between SBCC exposure and bed net use will be reduced in the second model. This reduced effect size will be considered as the direct effect of SBCC on bed net use. The indirect effect is the size of the effect of the mediating variable on bed net use.

Using this approach, we see that belief that mosquitoes are the sole cause of malaria explains 13 percent of the total effect of SBCC exposure on bed net use. Ideally we would want to include all the mediators in the model and the direct effect of SBCC would reduce to zero or near zero.

This analysis suggests that program implementers may want to consider messages that target the belief of what causes malaria in order to increase net use.

Taken together the approaches described in Parts 3 and 4 of this lecture represent a comprehensive framework for conducting outcome evaluations of SBCC interventions. The propensity score analysis described in Part 3 helps to answer the question: Did the SBCC intervention work? The mediation analysis presented in Part 4 of the lecture helps to answer the question: How did the SBCC messages work to change behavior? Combining these two approaches provide a complete outcome evaluation of SBCC interventions.

That’s the end of this lecture. To summarize let me highlight key points I hope you will take with you.

First, remember that outcome evaluations are learning exercises and you ideally want to learn about two primary questions: Was the program effective and how was the program effective? If you are able to answer these questions, you will be well equipped to offer guidance to program managers and implementers on the next steps that they should take.

Second, remember that there are several approaches that can be used to evaluate SBCC interventions. None of these are perfect and all require specific types of data and make specific assumptions. When deciding on the design for your outcome evaluation you should consider this range of approaches in light of the data that you were able to collect, the technical resources you have, and the strength of the evidence that you need to document the effects of the intervention. We believe that propensity score matching and mediation analysis offer the strongest evidence, but they also require a large survey and sufficient technical resources to do the analysis. For smaller interventions, it may make more sense to use one or more of the standard approaches described earlier in the talk.

Finally, here are some resources that may help guide you in your evaluation work. The paper by Kincaid and Do provides a more detailed description of the combined propensity score/mediation analysis approach for SBCC evaluation. The World Bank manuscript describes several highly technical approaches for outcome evaluation and offers alternatives other than propensity score matching. The Hornik chapter provides a nice summary of the standard approaches for SBCC outcome evaluations.

Well, thank you for spending some time listening to this talk. Best of luck in your evaluation work.

Resources

- Multivariate causal attribution and cost-effectiveness of a national mass media campaign in the Philippines. Kincaid DL, Do MP. J Health Commun 11 Suppl 2, 69-90 (2006).
- Handbook on Impact Evaluation: Quantitative Methods and Practices. Khandker SR, Koolwal GB, Samad HA, eds. World Bank: Washington, DC (2010).
- Evaluation design for public health communication programs. Hornik RC. In: Hornik RC, ed. Public Health Communication. Lawrence Erlbaum Associates: Mahwah, NJ; 385-408 (2002).

Speaker biography

Dr. Boulay is an assistant professor in the Department of Health, Behavior and Society at the Johns Hopkins Bloomberg School of Public Health. He has over ten years of experience monitoring and evaluating both social and behavior change communication and advocacy projects in Africa, Asia, Eastern Europe and in the United States and has worked on a range of health issues, including maternal and child health, family planning, reproductive health, HIV/AIDS and malaria. His areas of expertise include the use of statistical and econometric approaches in program evaluation and the use of social network analysis to examine the diffusion of program messages through family and peer networks. He teaches an advanced graduate course on social network analysis at the Bloomberg School of Public Health.

Questions? Contact gabrielle.hunter@jhu.edu

To access the full course content, including slides and videos, go to [SBC Learning Central]. [Getting Started (PDF)].

Acknowledgments

This presentation is made possible by the generous support of the American people through the U.S. Agency for International Development. The contents are the responsibility of the presenter and do not necessarily reflect the views of USAID or the U.S. government.