

To be published as a WHO Monograph in early 2000.

---

Title:

**Issues in Evaluating Mass Media-Based Health Communication Campaigns**

---

Authors:

*Vicki Freimuth, Ph.D.*  
Associate Director for Communication  
Office of the Director  
Centers for Disease Control and Prevention

*Galen Cole, Ph.D., M.P.H.*  
Senior Scientist  
Division of Health Communication  
Office of Communication  
Office of the Director  
Centers for Disease Control and Prevention

*Susan Kirby, Dr.PH.*  
Senior Scientist  
Division of Health Communication  
Office of Communication  
Office of the Director  
Centers for Disease Control and Prevention

---

Contact Person:

*Galen E. Cole, PhD, MPH*  
Senior Evaluation Scientist  
Division of Health Communication  
Office of the Director  
Centers for Disease Control and Prevention  
1600 Clifton Road, Mailstop D-42  
Atlanta, GA 30333  
Tel: (404) 639-7275; Fax: (404) 639-7391  
E-mail: gxc9@cdc.gov

---

Acknowledgments:

Carole Craft and Linda Cockrill for editing the paper

---

## INTRODUCTION

Most premature deaths in developed countries can be linked to action or the lack of action by individuals and/or communities (1). As a result, public health practitioners have developed interventions to promote healthful attitudes and actions and to suppress those which place life and health in jeopardy. Health communication, which we define as *the study and use of strategies to inform and influence individual and community decisions that enhance health*, plays an increasingly central role in these interventions.

Communication may be a dominant player or may have a supporting role in an intervention. Some roles may include communication strategies such as public relations, where the objective is to get the health issue on the public agenda; entertainment education, where desired behaviors are modeled in an entertainment program; and media advocacy which entails using the media as an advocacy tool to achieve policy level change. All of these strategies may include a range of communication activities that can occur at the individual, small group or mass media level. This paper addresses only those communication activities that use mass media outlets and, more specifically, the issues surrounding the evaluation of the development, implementation, and effects of mass media health communication campaigns. These evaluation issues will be discussed under the commonly known headings of formative, process and summative evaluation.

## HEALTH COMMUNICATION CAMPAIGNS

Rogers and Storey (2) maintain that health communication campaigns have four defining characteristics. These campaigns strive to (1) generate specific outcomes or effects (2) in a relatively large number of individuals (3) usually within a specified period of time and (4) through an organized set of communication activities. Health communication campaigns that rely on mass media outlets frequently consist of a series of television and radio public service announcements (PSAs) or paid commercials with collateral print materials such as posters, booklets, and brochures.

Most health communicators would agree that there are a common set of variables considered in the development of a mass media health communication campaign and a common set of outcomes that one can reasonably expect as a result of a communication experience. Communication development or

“independent variables” can be categorized into four broad areas: 1) psychosocial attributes of the receiver, 2) the source or spokesperson, 3) settings, channels, activities, and materials used to disseminate the message, and 4) the message itself, including content, tone, type of appeal, audio characteristics, and visual attributes. Taken together, any combination of these four independent variables constitutes what we refer to as the Communication Strategy (3, 4, 5, 6, 7, 8, 9, 10). The outcomes or “dependent variables” of a mass health communication effort may be categorized into six broad areas which include: 1) exposure, 2) attention, 3) comprehension, 4) yielding, 5) attitude change, and 6) behavior (4, 5, 11, 12). We acknowledge at the outset that these outcomes are not exhaustive, nor do we mean to imply that the progression of these outcomes are linear in nature. We do, however, believe that these terms will provide a common language pertinent to this discussion.

### **FORMATIVE RESEARCH AND EVALUATION ISSUES**

The research carried out prior to the implementation of a mass media-based health communication campaign is often referred to as formative research or formative evaluation. This “pre-implementation research” assists in understanding and developing effective communication strategies and tactics aimed at mitigating or eliminating problems (6, 9, 13, 14, 15).

Our purpose in this section is to discuss issues related to the formative research and formative evaluations carried out during the developmental stages of mass communication strategies and tactics. These issues pertain to 1) the data required to understand and profile the receiver characteristics of audiences that are the target of mass health communication, and 2) the evaluation or “pre-testing” of communication strategies and tactics prior to their implementation.

#### ***Data Issues***

The strategic development of a mass-mediated health communication campaign requires descriptive and analytic epidemiologic data to understand the nature and extent of the health problem as a basis for determining 1) whether mass communication is an appropriate intervention, 2) which audience(s) are the most appropriate targets of a mass communication intervention, and 3) what the overall goal of the communication should be. In addition to relying on traditional epidemiologic data, health communicators also need data required to segment and characterize potential audiences on independent variables that

have the most bearing on how one communicates with them [\(6, 9, 13, 14\)](#). As stated above, this includes data on the 1) psychosocial attributes of the receiver, 2) source or spokesperson, 3) settings, channel-specific communication activities, and materials that are used to support communication activities, and 4) the message itself, including content, tone, type of appeal, audio characteristics, and visual attributes. These data allow the communicator to disaggregate the population of interest into homogeneous subgroups or audience segments. Health-related audience segments are usually defined by being alike in one of two ways; a) regarding predictors of the behavior (similar levels of self-efficacy, social norms, or knowledge) or b) regarding communication strategy factors (e.g., they are motivated by a fear based message, or they prefer a lay person to communicate the message).

Although there are numerous sources of health-related data as well as many sources of data on consumers used for marketing purposes, to our knowledge the only national data generated in the U.S. which combines health behavior predictor data with data on communication variables is called *Healthstyles*® [\(14\)](#) and has been collected since 1995 by Porter Novelli, a social marketing firm located in Washington D.C. The lack of data puts communication planners in a position of having to 1) collect primary data, 2) merge or retrofit epidemiologic and marketing data, or 3) plan interventions without a clear understanding of communication-relevant differences that may exist in the populations they are targeting. Obviously, if adequate time and resources are available, the first option is preferred. However, when working on a short timeline with a limited budget it would be helpful to have multi-variate datasets [\(16, 17, 18, 19, 20, 21, 22, 23, 24\)](#) available that provide information on salient health-related and communication variables. In view of this, we recommend that researchers take steps to systematically link or create databases that provide the etiologic data required to understand health behavior incidence and prevalence [e.g., Behavioral Risk Factor Surveillance Study [\(25\)](#)], data that helps us understand what is driving health problems that can be addressed by mass communication interventions, and the communication data that can help planners understand how to effectively tailor mass communication strategies to the receiver characteristics of homogenous segments of the population.

### ***Pretesting Communication Strategies and Tactics***

Once an audience is segmented into groups who share similar characteristics that are important to the communication process, specific communication strategies and tactics can be crafted for each

segment of the population. The crafting of a communication strategy that is tailored to the health information and communication needs of a particular audience segment is facilitated by providing those responsible for developing the strategy with a creative brief which consists of a profile of the health-relevant knowledge, attitudes, actions and communication-related characteristics of each target audience. Although, at this point in the communication planning process, existing research may provide much guidance on each variable in the communication strategy, research on how to put those variables together most effectively for a particular audience segment is rare. Hence, health communication planners rely on a type of formative evaluation referred to as pre-testing.

In short, pre-testing is a process for systematically determining which combination of options represented by each communication variable (i.e., the communication strategy) tend to be most effective in achieving the communication objectives. This type of formative research shares characteristics of both process and summative research in that it can be designed to examine both the simulated delivery and the effects of a communication strategy and its tactics. At the same time, pre-testing is different from process or summative evaluation research in that it is carried out before final production and execution of a communication strategy to determine whether each element in the mix helps achieve the communication objectives of the project and meets the information needs of the intended audience (27, 28). A point which will be made under the summative evaluation section of this paper is relevant here. That is, because it is so difficult to directly attribute changes in individuals to a mass communication intervention, a high priority should be placed on pre-testing a strategy before it is executed to ensure that it is feasible, it produces intended cognitive effects in a sample of individuals who are representative of each target audience, and it does no harm.

While it is considered an indispensable formative evaluation method, there are some issues surrounding how pre-testing is carried out. Namely, rigor of the research methods employed and comparison data for decision-making. This is particularly true with focus group interviews which have steadily increased in the non-profit and health community (29). A simple search of *Medline*® indicated a 266 percent growth over eight years in reported focus group studies from 45 (1988-1991) to 165 (1992-1995). Although all research methods have inherent advantages and disadvantages, the problems posed by the potentially inappropriate use of focus groups are worthy of a brief discussion.

Focus groups are based on conducting a series of small group discussions with members of

intended audience segment. A series of groups is recommended because the unit of analysis is the group itself, not each member of the group. Because of various constraints, program planners often conduct too few groups for each segment of interest. This tendency can result in conducting one focus group session for each segment type (e.g., one group each of black females, white females, and Asian females), which is not adequate for drawing research conclusions and is far too few to find any between-group differences (29). Another problem in using focus group research is the temptation to quantify participants' answers (e.g., by asking for a hand count on agreement) which leads others to believe that the data may adhere to the rules for quantitative data integrity, such as independent observations or central limit theorem. Focus group authors have long cautioned practitioners not to quantify results (29, 30, 31) because it misleads readers and destroys the true value of qualitative research, which is to gain a richer and deeper understanding of a topic, not a more precise or accurate measurement. Health communication planners need to be able to judiciously use focus groups to their best advantage while maintaining a high level of confidence in the findings. Focus groups, not to be confused with group interviewing, should be not be used for message pretesting, except to explore answers to quantitative measures.

There are several other qualitative research methods that can be used along with quantitative methods to gather information in connection with pre-testing messages. These include methods such as case studies, one-on-one interviews, and record abstraction (32, 33, 34). In most instances, we prefer one-on-one interviews or central intercept interviews, as they are often called in communication research, to test messages. We prefer this method of message testing because: 1) it's easier to connect with harder-to-reach respondents in locations convenient and comfortable for them; 2) we can access an increased number of respondents within the intended population if an appropriate location is selected; 3) it's a cost-effective means of gathering data in a relatively short time; 4) we can get a larger sample size than focus groups, and 4) these one-on-one interviews tend to eliminate group bias that is possible in focus groups.

Lastly, once formative researchers have some pretest data in hand, little comparison data are available to help decide if the pretested materials performed "well enough" to create change in a real world setting. Aside from the Health Message Testing Service (35), few health communication programs have conducted quantitative message pretesting, published their findings, or related pretesting data to

outcome evaluation data. Without knowledge of pretesting data and actual communication outcomes, health communication planners cannot forecast how well a communication strategy will help reach communication objectives. The advertising world, which refers to message pretesting as copytesting, may have some useful models that can be adopted to help overcome this lack of data.

Most advertising agencies employ some method of copytesting [\(36\)](#) and have established marketing surveillance systems for the purpose of consistently collecting, analyzing, and cataloging the data generated by the copytesting process. The rigor and systematic collection of these data is demonstrated by the 1982 Positioning Advertising Copytesting (PACT) agreement [\(37\)](#). The document outlining the PACT agreement, prepared by 21 of the leading advertising agencies in 1982 [\(37\)](#), articulates nine principles of copytesting which target multiple measures, representative samples, reliability, and validity. Since the PACT agreement was published in 1982 [\(37\)](#), a plethora of research firms have been established to help deliver on these nine principles (BehaviorScan, the Starch Report, AHF Marketing Research, ASI Marketing Research, Gallup and Robinson, McCollum/Spielman, and Mapes and Ross [\(38\)](#)). These firms research, track, and collect copytesting data for primary purchasing and reselling to retail organizations. These data help identify the most effective and efficient communication strategy for the marketing communication dollar.

While it is true that outcomes like consumer recall and purchase data are easily assessed and collected in the retail marketing world, cognitive and behavioral outcomes of health communication activities are not as easily assessed or collected. Thus, health communication campaigns that rely on mass media outlets are usually challenged with making formative decisions based on relatively little data and, even in the best situations, making decisions without up-to-date information or comparison data. However, as was noted above, for health communication planners to bring the best messages to the prevention marketplace to attack the root causes of the health problems of our times, they will need timely comparison data systems, not unlike those established by the private sector, to identify and improve weak and inadequate programs before implementation.

### **PROCESS EVALUATION ISSUES**

Generally speaking, *process evaluation* is used to answer questions about whether a program is delivered as planned [\(28\)](#). In the present context, process evaluation addresses questions concerning

how well and under what conditions a mass media health communication campaign was implemented, and the size of the audience that is exposed to the message. A number of issues should be understood and addressed when planning and conducting a process evaluation of mass media campaigns. These include: 1) the utility of process evaluation, 2) theoretical considerations and cause/effect attributions, and 3) changing an intervention during the course of an evaluation.

### ***The Utility of Process Evaluation***

The clamor for data on the intended effects of campaigns by stakeholders has led many evaluators to become preoccupied with program impacts and outcomes (26, 39, 40). This focus on the effects of interventions has led many evaluators to rely heavily on controlled experimental methods. An enticing feature of these methods is that an understanding of how a campaign works is not necessary to estimate its net effects through random experimental methods (41). Hence, evaluators can satisfy the demand for effects without carefully considering, through the auspices of process evaluation, the program mechanisms that produce these effects.

The negative result of evaluating outcomes without knowledge of implementation is that stakeholders receive very little information upon which to act (26). That is, even though an experimentally based evaluation may demonstrate that an intervention produces intended effects, if the implementation processes of a campaign are not accounted for through a formal process evaluation, there is very little basis for taking action to improve a program because stakeholders lack information about what produced the observed outcomes (26). For example, if an evaluation of a PSA campaign to increase moderate physical activity among adults in a particular community does not include surveillance to determine what proportion of the target audience is exposed to the PSA, it is impossible to attribute observed effects to the intervention. Weiss (42) makes this point wherein she says: *"Does it make any difference...whether the program is using rote drill, psychoanalysis, or black magic? There are evaluators who are sympathetic to such an approach. They see the program as a 'black box,' the contents of which do not concern them; they are charged with discovering effects. But, if the evaluator has no idea of what the program really is, s/he may fail to ask the right questions"* (in 26). The point is, without process evaluation, one cannot differentiate between a bad campaign and one that is poorly implemented. This is

particularly true if one is trying to improve campaign effects through modifying, enhancing and, if necessary, eliminating campaign processes. In sum, the best evaluation considers both processes and effects.

### ***Theoretical Considerations and Cause/Effect Attributions***

Although many recognize the importance of both process and summative evaluation, these assessments are sometimes conducted independently as if there is no connection between the two. This results in a post hoc “cut-and-paste” job where, after the data are collected on both process and outcome markers, the evaluators attempt to link effects with specific processes. To prevent this, we advocate that health communication evaluators clearly delineate, a priori, linkages between program processes and intended outcomes. The importance of this is explained by Patton (26), Weiss (40), and Chen (41) who all state that evaluators should define a campaign’s “theory of action” before initiating evaluation. That is, before beginning the evaluation, each important intervention process (independent variable) should be explicitly linked with each desired outcome (dependent variable). This approach is often referred to as theory-based evaluation (41).

In theory-based evaluation, the standard for comparison is the program’s theory, or sub-theories, if the evaluation is aimed at examining sub-components of the program. Therefore, the first phase in theory-based evaluation is theory construction. This requires an understanding of what program theories are and how best to develop them.

If the program theory is examined as a whole, only one program theory is necessary. However, if one wants to know about different sub-components -- domains of the program theory, as Chen (41) refers to them -- such as its development, delivery, cost, or effects domain, a separate sub-theory must be constructed for each program domain to be evaluated. Chen (41) notes that program theory domains can be considered independently (basic types) or in some combination (composite types). In short, the evaluator must construct a separate theory for each basic and/or composite domain selected for a theory-based evaluation, with each theory serving as a standard of comparison.

The idea of comparing what theoretically should happen to what actually happens, in terms of the performance of the program, and/or comparing a problem theory against the reality of the problem as discovered by a theory-based evaluation is a rather simple notion. What is not so straightforward is *how to*

construct a problem theory or an expected program theory of action (i.e., the standard of comparison) that accurately reflects how a program is supposed to perform and the nature of the problem(s) it is designed to overcome. Fortunately, a number of different strategies have been developed to assist with this process (40, 43, 44). These strategies help users systematically construct program development, implementation, and cause/effect theories and sub-theories that serve as standards of comparison against which evaluation data can be compared to identify the extent of discrepancy or congruence that exists between how the campaign activities are supposed to bring about intended effects and what actually happens (44).

### ***Changing an Intervention During the Evaluation***

Evaluation is an iterative process designed to provide relevant and timely feedback to stakeholders to make decisions aimed at improving the program. The implicit assumption is that if this feedback dictates the need to change the program to improve it, the program should be changed, particularly with social marketing programs where the aim is to respond rapidly to feedback. However, changing a program is at odds with the scientific dictum to standardize or “keep the intervention constant” throughout the course of an evaluation.

To overcome these conflicting purposes we suggest that evaluators and program implementers agree, from the outset, to an appropriate schedule specifying when feedback will be reported and, if appropriate, changes in the program will be made. Threats to the validity of the findings can be minimized by ensuring that 1) changes in the campaign processes are documented, 2) process evaluation tracking protocols are modified to account for these changes; and 3) measurements are taken on key outcome variables both before and after important changes are made in the implementation process. This will ensure the constancy needed to pick up effects that may result from changes in the campaign processes while allowing planners to respond to timely and relevant feedback that can be used to improve the program.

## **SUMMATIVE EVALUATION ISSUES**

Summative evaluation of a mass-mediated health communication program assesses whether the

intended audience was reached and the impact and outcome objectives of the program were achieved to the satisfaction of the stakeholders. In this section we will discuss issues around both of these types of summative assessment.

### ***Issues Pertaining to the Summative Assessment of Reach***

In order for a message to have a desired influence, receivers must first attend to it (5). Hence, an early effect of communication that must be observed in a summative evaluation is whether the intended audience paid attention to the desired message. As stated previously, even if a program is implemented as planned, and desired effects result, these effects can not be attributed to the intervention unless there is evidence that the campaign actually reached the intended audience.

A necessary first step in determining whether the intended audience was reached by mass communication messages is to determine whether a message airs and the number of times it airs. If one can afford paid advertising, this process is greatly simplified because the time, place, and frequency of airing can be controlled. However, for a number of reasons--with cost being a leading factor--paid advertising is seldom used by public sector health communicators in the United States.

When paid advertising is not an option in a mass media campaign, health communicators often rely on PSAs which are aired at no cost to the producer. Unfortunately, airing of PSAs in the U.S. is at the discretion of Public Service Directors at the various television and radio stations, making the tracking of airings difficult. Attempts to overcome these difficulties have included relying on services that monitor commercials and PSAs. For example, Nielsen Media Sigma Service (NMSS) operates an electronic tracking service that detects the airing of PSAs in over 1,100 broadcast stations (including 40 Spanish language stations) in all 211 designated market areas plus 28 national cable networks. This service ascertains the number of times a PSA plays, the market(s) where it played, the station call letters, air date, and air time. This monitoring goes on 24 hours a day, seven days a week.

To get some idea concerning the extended reach of broadcast and print media that may have been triggered by our mass media campaigns, we consistently track both broadcast and print media. The services used to track broadcast media is called *Video Monitoring Services, Inc. (VMS)*. The print news tracking service we use is *Lexis-Nexis*.

VMS monitors news and public affairs programming in 46 of the top media markets. This includes

300 local television stations and 50 network and cable channels such as CNN, CNBC and MS-NBC. VMS also monitors selected news radio programming generated by 60 radio stations in 15 of the top media markets.

Lexis-Nexis continually updates and maintains a Regional News library which consists of a combination of news sources grouped together by geographical area. It contains more than 125 full-text U.S. regional news sources together with selected documents from Business Dateline, ABI/INFORM, and abstracts from Miami Herald and Philadelphia Inquirer. The UPI State & Regional wire service is also included.

Together, NMSS, VMS and Lexis-Nexis services allow us to estimate the overall reach of both broadcast and print media. Although these data satisfy the need to determine whether, when, and where a PSA is aired, along with the extended reach of collateral media that may have been triggered by the campaign, these services and the data they generate do not account for who attended to, comprehended, and yielded to the key messages of a campaign. Further audience research must be carried out to make this determination.

In the U.S., to determine who was watching or whether those who were watching were attending to the central messages broadcast on television as a part of a national health communication campaign one can rely on services like the Nielson Station Index (NSI) which generates information regarding the TV viewing behavior of individuals (>100,000) living in randomly selected households in each of the U.S. TV markets. NSI characterizes viewers demographically by their age and gender. Data are collected using diaries for each TV in a participating home. Participants record the programs they watch and for how long, the station the program was aired on, and the date and time the program was aired. Data collected with diaries are further verified and adjusted based on TV "set meters" that electronically capture household viewing events in a sample of TV markets.

To further characterize the audiences who "may have" viewed a particular message, NSI data can be merged with geo-psychographic data aggregated into neighborhood clusters that represent demographic and/or psychographic profiles of individuals living in different neighborhoods in various locations across the U.S. Merging these data with NSI data allows for the indirect approximation of the psychographic characteristics of those who view a message in question. For example, through their PRIZM cluster analysis system the Claritas Corporation--perhaps the most prominent vendor of geo-

psychographic data clusters--provides information on households categorized to one of 62 neighborhood audience segments based on six criterion factors: social rank, household composition, residential mobility, ethnicity, urbanization, and types of housing. Also available in the database is information on media habits, small and large purchase patterns, political beliefs, geographic location, and demographics. The point here is that the process of merging a variety of data sets allows for an indirect approximation of who is watching what and when they are watching.

What is still missing, however, is whether these audiences attended to the messages. Some approaches to determining whether a particular audience attended to and comprehended messages are to: 1) conduct a general population survey to determine audience awareness of a campaign; 2) add specific relevant questions to an Omnibus survey; 3) rely on data collected in national probability sample surveys; and/or 4) add tags to a televised message which are designed to motivate viewers to call a particular number for more information with the assumption that a burst of calls just after the airing of a the message with such a tag almost certainly indicates the audience attended to the messages. Questions directed at those who call in can help further determine whether those who attended to the message actually understood it. All of these summative evaluation approaches have been used at the CDC in attempts to monitor the reach of our HIV/AIDS health communication efforts carried out by what was the National AIDS Information and Education Program (45).

As with tracking electronic media, one must be highly creative in determining who is exposed to messages in newspapers or magazines and even more so with collateral materials such as brochures, flyers, posters, and billboards. This often becomes labor intensive and expensive.

### ***Issues Around Assessing Intended Effects***

Flay and Cook (46) have identified three models which have been used to conduct summative evaluations of health communication program effects. These are the advertising model, the impact-monitoring model, and the experimental model. The advertising model is used most frequently and consists of a baseline survey before the program is implemented followed by another survey at the end of the program. The evaluation of the Cancer Prevention Awareness Campaign (47) is a representative example of this approach. A national probability survey was conducted before the launch of the campaign and again a year later, after a multi-channel cancer prevention campaign was implemented. Materials

included booklets, radio and TV PSAs, and special events. The evaluation compared knowledge of risk factors and concern about cancer before and after the campaign. This evaluation model is simple and often criticized because the lack of a control group prohibits establishing a direct cause and effect relationship between the campaign and its outcomes.

The impact-monitoring model uses routinely collected data from a management information system to monitor outcomes and impacts of a health communication campaign. For example, as part of their evaluation of the national AIDS campaign, CDC examined knowledge, attitude, and behavior measures from its annual National Health Interview Surveys. This method is easy and cost-effective, but it usually measures only behavioral outcomes and often fails to provide information which can explain successes and failures.

The experimental model contrasts two or more equivalent groups, one of which is a no-treatment control group. An anti-smoking campaign, designed to recruit women cigarette smokers with young children to call for information on quitting, used this evaluation model [\(48\)](#). The campaign included a mix of professionally produced broadcast and print media which encouraged mothers who were smokers to call the National Cancer Institute's Cancer Information Service (CIS) for information on quitting. Careful placement of media messages was possible because paid advertising was used. Fourteen media markets in New York, Pennsylvania, and Delaware were size-matched and one of each pair was randomly assigned to the experimental group who received the campaign and the other to the control group. Response to the campaign was gauged by monitoring calls to the area CIS offices from smokers residing in these experimental and control media markets. This model is usually considered the most rigorous, but has been challenged as inappropriate for evaluating what is essentially a messy social process [\(49\)](#).

The choice of an appropriate model of evaluation depends on an understanding of the way health communication campaigns work. Hornick [\(49\)](#) presented a compelling argument against the controlled randomized experimental design. He contrasted the limited effects attributed to such well-known community-based health promotion efforts as Stanford's Three and Five City Studies, the Minnesota and Pawtucket Heart Health Programs, and the Community Intervention Trial for Smoking Cessation (COMMIT) with the impressive evidence of behavioral change from the National High Blood Pressure Education and Control Program (NHBPEP), the original televised smoking counter-advertising campaign

between 1967 and 1970, the public communication around the AIDS epidemic in the U.S., and the current California anti-smoking campaign. He attributed this surprising contrast in effectiveness to the constraints imposed by the research design itself. It is quite misleading to think that no background communication on a health issue is occurring in control communities; treatment communities may only have slightly more exposure to messages about these issues. Stanford, for example, claimed that it provided 25 hours of exposure on average to heart disease messages over five years in its treatment communities. This estimate suggests that most people only received one hour of messages per year on each of the five behaviors promoted. Hornick (49) contrasted this limited exposure to the more intense scale of the NHBPEP which represents the complex social diffusion process--deliberate communication messages, the conversations that ensue, the coverage by other media sources, the demands put on institutions which then respond, health institutions which offer different advice and treatments, the commercial institutions which make new products and advertise different benefits, and on the political institutions which change public policy to be supportive of the health behaviors. He argued that communication is a social process, not a pill, and should be evaluated as such.

Hornick's reasoning also reinforces the difficulty in disentangling communication effects from those of other intervention components or disentangling the effects of several communication activities. If we assume that a complex social change process has occurred, we have to either develop more sophisticated tools for measuring this diffusion process and disentangling its separate components or be content with assessing overall effects without attribution to individual components of the intervention. It may be reasonable to expect practitioners to do only the latter in routine evaluations of campaigns but to ask health communication researchers to design studies to capture this complex social diffusion process and discover how individual communication components contribute to it.

Above all, we must resist the effort to design rigorous, controlled experimental studies that strive to compare the effects of individual communication products such as pamphlets, PSAs, and posters with the goal of answering which product is the most effective to use across all situations. That kind of evaluation is inconsistent with the research and practice literature that recommends multiple messages and channels and cautions that finding the right channels to reach the right audiences with the right messages delivered by the right sources at the right times is best answered with formative research conducted early in the campaign development.

Even after an appropriate evaluation design is selected, summative evaluations of health communication campaigns face some serious methodological issues. The most common problems are described in the following sections.

### ***Measurement Problems***

Frequently, health communication components of interventions have several objectives. One of the most critical measurement problems involves determining which effects to measure. Earlier we described these potential effects, i.e., the individual has to attend to the message, comprehend it, relate it to other information he has, yield to it, and translate his new beliefs into behaviors that are then tested repeatedly. Does the evaluator measure comprehension, attitude change, or behavior change? One might argue that the further along this chain you measure, the more important the effects. On the other hand, the potential effects of the messages decrease as one measures further along this chain, as does ability to control extraneous variables. For example, is it realistic to assume that the direct cause of a smoker quitting is an anti-smoking PSA? Undoubtedly, the causal chain is more complicated than that. In addition, many behavioral changes advocated in health messages are impossible to observe directly. For example, how can hypertensives' use of medication or a woman's breast self-examination be observed? In such cases, self-reported behavior is measured. Such measurement is subject to error because of the tendency to over-report socially desirable behavior. Some evaluations are able to validate self-report measures with behavioral or physiological data. For example, smoking cessation studies often validate a percentage of their self-report measures with saliva cotinine testing and proxy data from people who are willing to observe the smoking behavior of an individual who participates in the cessation program.

Most summative evaluations of health communication programs attempt to measure exposure to the messages by questioning respondents about their recall of the messages. Unaided recall generally produces an artificially low estimate of audience exposure. Most evaluators use some form of aided recall; that is, they provide the respondent with some information about the message and then ask if the respondent remembers hearing or seeing it. With the use of aided recall, however, there will be some over-reporting of exposure. Over-reporting occurs when respondents acquiesce or try to be helpful by giving what they think is the desired answer. In an attempt to avoid over-reporting, verbatim descriptions of the messages often are requested. Only respondents whose descriptions can be clearly tied to the

message in question are identified as having been exposed. This approach requires rigorous coding procedures to classify the respondents.

A measurement compromise we suggest is the use of unaided questions followed by a series of aided questions, such as, "Recently, ads showing fatal car accidents in which seat belts were not worn have been broadcast. Have you seen any of these ads?" Estimates of the magnitude of error due to over-reporting can be calculated based on measure of (spurious) reported awareness among respondents not exposed to the ads (i.e., in a control condition) or on reported awareness to bogus messages.

### ***Sampling Problems***

As we described in an earlier section of the paper, most mass media health communication campaigns are targeted to a specific segment of the audience but the evaluation frequently is not limited to that segment. When sampling for a post campaign evaluation, how does one find women who have not had a mammogram, individuals who build campfires in the forest, hypertensives who do not take their medication regularly, or drivers who do not wear seat belts? At best, there are some demographic data available, but these are far from perfectly descriptive of the target group. This difficulty in identifying the target group is compounded by the frequently low exposure to many of these messages. It is not uncommon for recall of a health message to be as low as 10 percent. Consequently, every random sample of 1,000 may only yield 100 persons who remember seeing the message. Imagine how these numbers decrease if we are looking for women who have not had a mammogram who recall seeing messages recommending mammograms. Most surveys used to measure the effectiveness of health messages need to screen respondents carefully, a time-consuming and costly process.

## **SUMMARY AND DISCUSSION**

Because of communication's varied roles in public health interventions, there are a number of issues involved in evaluating health communication campaigns. This paper has highlighted these issues and provided a number of ways to resolve them within the context of formative, process, and summative evaluation.

Formative evaluation of health communication campaigns involves conducting research to assist in developing the most effective communication strategy and then testing that mix to forecast how

effective it will be in reaching communication objectives. This entails breaking the audience into smaller homogeneous segments and then characterizing, or profiling, those segments in order to more closely tailor campaign messages and implementation. Profiling audience segments can best be accomplished with the benefit of datasets that include both etiologic data on the distribution and determinants of health problems that may be mitigated by mass communication, and information on variables that allow planners to understand how to best communicate with each audience targeted by the communication strategy (50).

Formative evaluation carried out prior to the implementation of a communication strategy is key to ensuring that the strategy is feasible, produces intended effects in each target audience, and does no harm. While it is considered an indispensable formative evaluation method, there are some issues surrounding how pre-testing is carried out. Namely, which research methods should be employed and how they are used in the pre-testing process. In this vein, we recommend that formative evaluators judiciously select quantitative and qualitative methods that are best suited to pre-testing and that these methods are employed in a technically acceptable manner.

Once a mass media health communication campaign is underway, process evaluation begins to assess how the program is rolling out and working. Without knowing how the campaign worked, we cannot determine whether the program brought about desired effects or if other factors influenced those effects. This knowledge is also critical in determining what aspects of a mass media campaign should be changed or eliminated, if any, to improve the campaign. We recommend applying the principles of theory-based evaluation to construct models that explicitly state how the program will bring about intended effects in order to have a basis for comparing how the program actually worked. Furthermore, we recommend that health communication researchers design studies to capture the complex social diffusion process that occurs in and around a mass media campaign in an effort to systematically discover how individual communication components contribute to the campaign process as a basis for explaining and replicating them when they work.

Process evaluation can be used to determine a program's effectiveness while it is ongoing. This allows for changes to be made to a program midstream in order to increase the likelihood of desired outcomes. Implementers and evaluators of the program should agree from the outset on a schedule for reporting feedback and making informed changes in the campaign to assure its maximum relevance, efficiency and effectiveness.

Summative evaluation of a mass media health communication campaign aims to determine whether the intended audience was reached and the objectives of the program were achieved. This type of evaluation, however, is complicated by several factors.

First, while we can monitor whether messages were disseminated, it is more difficult to assess whether the intended audience was exposed to the messages and attended to them. In the case of PSAs in a national media campaign, for example, a broadcast play verification company can help determine if the intended audience was exposed to the campaign's messages. But to find out if the audience attended to those messages, it is often necessary to conduct surveys, rely on data collected in national probability sample surveys, or use tags on PSAs that are designed to motivate viewers to call a particular number for more information. The latter method assumes that a burst of calls just after a PSA's airing will indicate the audience saw and paid attention to the PSA.

To determine whether the campaign messages had the intended effect(s), we can employ one of three models—advertising, impact-modeling, and experimental. The advertising model, consisting of a baseline survey before the campaign's implementation and another after its conclusion, is most frequently used, but it is also draws criticism because it lacks a control group. The impact-monitoring model uses routinely collected data from a management information system to monitor behavioral outcomes and impacts of the campaign. This method, while easy and cost-effective, fails to provide information that explains success or failure. The experimental model contrasts two or more equivalent groups, one of which is a no-treatment control group. We argue that this method is imperfect because it assumes that no background communication is going on in control communities, a belief that is unrealistic.

Even after a summative evaluation design is selected, a number of concerns arise. When using surveys, over-reporting can result, especially when dealing with socially desirable behaviors. For example, will drivers who say they saw PSAs on safety belts admit that they do not use them? There is also difficulty in identifying the target audience after the campaign. Exposure may be low, many people may have forgotten seeing a message, and people may not admit to seeing a message if they have not adopted the behavior encouraged in the message.

We have identified and discussed issues that are important to consider in the conduct of formative, process, and summative evaluations of mass media health communication campaigns. These issues and our recommendations pertaining how they might be resolved should provide a basis for further

improvements in conceptualizing, planning, implementing, and reporting feedback on evaluations aimed at improving mass media health communication campaigns to promote health-enhancing behaviors.

## REFERENCES

1. McGinnis, J.M., & Foege, W.H. Actual causes of death in the United States. Journal of the American medical association, 270:18, 2207-2212 (1993).
2. Rogers, E. M. & Storey, J. D. Communication campaigns. In: Berger, C. & Chaffee, S., eds. Handbook of communication science. Sage Publications, Newbury Park, California, United States, 1987.
3. Flora, J. A., & Maibach, E. W. Cognitive responses to AIDS information: The effects of issue involvement and message appeal. Communication research, 759-774, (1990, December).
4. Flay, B.R., DiTecco, D., & Schlegel, R.P. Mass media in health promotion: An analysis using extended information-processing model. Health education quarterly 7(2):127-147 (1980).
5. McGuire, W. J. (1989). Theoretical foundations of campaigns. In: Rice, R. E., & Atkin, C.K., eds. Public communications campaigns (2nd ed., pp. 43-65). Sage Publications, Newbury Park, California, United States, 1989.
6. Sutton, S. M., Balch, G., & Lefebvre, C. Strategic Questions For Consumer-Based Health Communications. In: 5 a day for better health: NCI media campaign strategy (pp. 1-12). Washington D.C.: National Cancer Institute, 1993.
7. Gorn, G.J. The effects of music in advertising on choice behavior: A classical conditioning approach. Journal of marketing, 46: 94-101 (1982).
8. Messaris, P. Visual persuasion, Sage Publications, Thousand Oaks, California, United States, 1997.
9. Donohew, L., Lorch, E., & Palmgreen, P. Sensation seeking and targeting of televised anti-drug PSAs. In: Donohew, L., Sypher, H., & Bukoski, W. eds. Persuasive communication and drug abuse prevention, Lawrence Earlbaum Associates, Hillsdale, New Jersey, United States, 1991, pp. 209-226.
10. Prochaska, J.O., Norcross, J.C., & DiClemente, C.C. Changing for good, Avon Books, Inc., New York, New York, United States, 1994, pp. 287-289.
11. Backer, T.E., Roger, E.M., and Sopory, P. (1992). Generalizations about health communication campaigns. In: Designing health communication campaigns: What works? Sage Publications, Newbury Park, California, United States, 1992, pp. 30-32.
12. Petty, R. E., Cacioppo, J. T., & Schumann, D. Central and peripheral routes to advertising effectiveness: The moderating role of involvement. Journal of consumer research, 1983, September, pp. 135-146.
13. Atkin, C. & Freimuth, V. (1989). Formative Evaluation Research in Campaign Design. In: Rice, R. & Atkin, C., eds. Public communication campaigns, 2<sup>nd</sup> ed., Sage Publications, Newbury Park, California, United States, 1989, pp. 131-150.
14. Maibach, E.W., Maxfield, A.M., Ladin, K., & Slater, M.D. Translating health psychology into effective health communication: The American health styles audience segmentation project. Journal of health psychology, 1(3): 261-277, (1996).
15. Slater, M.D. Theory and method in health audience segmentation. Journal of health communication, 1: 267-283, (1996).

16. Slater, M.D. Choosing audience segmentation strategies and methods for health communication. In: Maibach, E. & Parrott, R. L., eds. Designing health messages: Approaches from communication theory and public health practice, Sage Publications, Thousand Oaks, California, United States, 1995, pp. 169-185.
17. Slater, M. & Flora, J.A. Health Lifestyles: Audience segmentation analysis for public health interventions. Health education quarterly, 18(2): 221-233, (1991).
18. Patterson, R.E., Hanes, P.S., & Popkin, B.M. Health lifestyle patterns of U.S. adults. Preventive medicine, 23: 453-460, (1994).
19. Maibach, E. Psychobehavioral segmentation: Identifying audiences and tailoring cancer prevention programs. Presentation to the Harvard Center for Cancer Prevention, Harvard School of Public Health, December 4, 1996.
20. Maibach E.W. & Cotton, D. Moving people to behavior change: A staged social cognitive approach to message design. In: Maibach, E.W. & Parrott, R. L., eds. Designing health messages: Approaches from communication theory and public health practice, Sage Publications, Thousand Oaks, California, United States, 1995, pp. 169-185.
21. Velicer, W.F., Hughes, S.L., Fava J.L., Prochaska, J.O., & DiClemente, C.C. An empirical typology of subjects within stage of change. Addictive behaviors, 20(3): 299-320, (1995).
22. Williams, J.E. & Flora, J.A. Health behavior segmentation and campaign planning to reduce cardiovascular disease risk among Hispanics. Health education quarterly, 22(1): 36-48, (1995).
23. Morris, L.A., Tabak, E.R. & Lins, N.J. A segmentation analysis of prescription drug intervention-seeking motives among the elderly. Journal of public relations and marketing, 11: 115-125, (1992).
24. Albrecht, T.L. & Bryant, C. Advances in segmentation modeling for health communication and social marketing campaigns. Journal of health communication, 1: 65-80, (1996).
25. Behavioral Risk Factor Surveillance Survey. State- and sex-specific prevalence of selected characteristics--behavioral risk factor surveillance system, 1994 and 1995. Morbidity and mortality weekly report. 46:(SS-3), August 1, 1997.
26. Patton, M.P. Utilization-focused evaluation. Sage Publications, Thousand Oaks, California, United States, 1986.
27. Freimuth, V. Developing public service advertisement for nonprofit contexts. In: Belk, R.W., ed. Advances in nonprofit marketing, JAI Press, Inc., Greenwich, Connecticut, United States, 1985, pp. 55-94.
28. Cole, G., Pogostin, C., Westover, B., Rios, N., & Collier, C. Addressing problems in evaluating health-relevant programs through a systematic planning and evaluation model. Risk: Issues in health, safety and environment, 6(1), 37-57, (1995).
29. Krueger, R.A. Focus groups: A practical guide for applied research, second edition. Sage Publications, Thousand Oaks, California, United States, 1994.
30. Stewart D.W., & Shamdasani, P.N. Focus groups: theory and practice. In: Applied social science research methods series, Volume 20. Sage Publications, Newbury Park, California, United States, 1990.

31. Templeton, J.F. The focus group (revised edition). Irwin Professional Publishing, Burr Ridge, Illinois, United States, 1994.
32. Crabtree, B. F. & Miller, W. L. Doing qualitative research: Research methods for primary care, Volume 3 (pp 3-28). Sage Publications Inc., Newbury Park, California, United States, 1992.
33. U.S. Department of Health and Human Services. Making health communication programs work. NIH Publication Number 92-1493, April, 1992.
34. Patton, M. Q. How to use qualitative methods in evaluation. Sage Publications. Newbury Park, California, United States, 1987.
35. U.S. Department of Health and Human Services. Pretesting in health communications: methods, examples, and resources for improving health messages and materials. NIH Publication Number 84-1493. January, 1984.
36. Shimp, T.A. Promotion management and Marketing Communications, 2<sup>nd</sup> ed. The Dryden Press, Chicago, Illinois, United States, 1990, pp. 428-432.
37. PACT Document. Journal of Advertizing. 11(4): 4-29, 1982.
38. Stewart, D. W. Furse, D. H., & Kozak, R. P. A Guide to Commercial Copytesting Services. In: Leigh, J. H. & Martin, C. R. Jr., eds. Current issues and research in advertising (Ann Arbor, MI: Division of Research, Graduate School of Business, University of Michigan). 1983, pp.1-44.
39. Gardner, D.E. Five evaluation frameworks. Journal of higher education, 48(5): 571-593, (1977).
40. Weiss, C.H. Nothing as practical as good theory: exploring theory-based evaluation for comprehensive community initiatives for children and families. In: Connell, J. P., Kubisch, A. C., Schorr, L. B. & Weiss, C. H., eds. New Approaches to evaluating community initiatives: Concepts, methods, and contexts. The Aspen Institute, Queenstown, Maryland, United States, 1995, pp. 65-92.
41. Chen, H. Theory-driven evaluations. Sage Publications, Thousand Oaks, California, United States, 1990.
42. Weiss, C.H. Evaluation research: Methods of assessing program effectiveness. Prentice-Hall, Englewood Cliffs, United States, 1972.
43. Shern, D.L., Trochum, W.M.K., & LaComb, C.A. The use of concept mapping for assessing fidelity of model transfer: an example from psychiatric rehabilitation. Evaluation and program planning, 18(2): 143-153, 1995.
44. Cole, G. E. Advancing the development and application of theory-based evaluation in the practice of public health. American Journal of Evaluation, 20(3): 453-470, 1999.
45. Nowak, G.J., & Siska M.J. Using research to inform campaign development and message design: Examples from the "America Responds to AIDS" Campaign. In: Maibach, E. W., & Parrott, R. L., eds. Designing health messages: approaches from communication theory and public health practice. Sage Publications, Thousand Oaks, California, United States, 1995, pp. 169-185.
46. Flay B.R., & Cook, T.D. Three models for summative evaluation of prevention campaigns with a mass media component. In Rice, R. and Atkin, C. (Eds.), Public communication campaigns, 2<sup>nd</sup> ed. Sage

Publications, Newbury Park, California, United States, 1989, pp. 175-196.

47. U.S. Department of Health and Human Services. Technical report on cancer prevention awareness survey wave II. Bethesda, MD: National Institutes of Health, 1986.

48. Cummings, K. M., Sciandra, R., Davis, S., & Rimer, B. Results of an antismoking campaign utilizing the Cancer Information Service. Monographs - National Cancer Institute, 14: 113-118, 1993.

49. Hornick, R. Public health communication: Making sense of contradictory evidence. Presentation at the Annenburg public policy center public health communication meeting, Washington, D.C., 1997.

50. Chervin, D.D., Nowak, G.N., & Cole, G.E. Using audience research in designing public health initiatives at the federal level. Social Marketing Quarterly. 5(3): 34-39, September 1999.