TBM

ORIGINAL RESEARCH

Comparing strategies to assess multiple behavior change in behavioral intervention studies

Bettina F Drake, PhD,[1] Lisa M Quintiliani, PhD,[2] Amy L Sapp, ScD,[5] Yi Li, PhD,[3] Amy E Harley, PhD,[4] Karen M Emmons, PhD,[5] Glorian Sorensen, PhD[5]

[1]Department of Surgery, Washington University School of Medicine, 660 S. Euclid AveCampus Box 8100, St. Louis, MO 63110, USA
[2]Department of General Internal Medicine, Boston University Medical Center, Boston, MA, USA
[3]Department of Biostatistics, Harvard School of Public Health, Boston, MA, USA
[4]School of Public Health, University of Wisconsin, Milwaukee, WI, USA
[5]Center for Community-Based Research, Department of Society, Human Development and Health, Harvard School of Public Health, Boston, MA, USA

Correspondence to: B F Drake drakeb@wustl.edu

Cite this as: *TBM* 2013;3:114–121
doi: 10.1007/s13142-013-0195-2

ABSTRACT
Alternatives to individual behavior change methods have been proposed; however, little has been done to investigate how these methods compare. To explore the four methods that quantify change in multiple risk behaviors targeting four common behaviors, we utilized data from two cluster-randomized, multiple behavior change trials conducted in two settings: small businesses and health centers. Methods used were (1) summative, (2) z-score, (3) optimal linear combination, and (4) impact score. In the small business study, methods 2 and 3 revealed similar outcomes. However, physical activity did not contribute to method 3. In the health centers study, similar results were found with each of the methods. Multivitamin intake contributed significantly more to each of the summary measures than other behaviors. Selection of methods to assess multiple behavior change in intervention trials must consider study design and the targeted population when determining the appropriate method/s to use.

KEYWORDS
Multiple behavior change, Intervention, Methods, Multiple risk factors

Implications

Practice: Multiple health behavior change interventions are effective in reducing risk behaviors for major chronic diseases and the effectiveness of these interventions can be measured in a number of ways.

Policy: Resources should be directed toward better measurement of health risk behaviors in order to more appropriately test and identify effective multiple health behavior change interventions.

Research: Research needs to be directed towards better understanding methods used to assess multiple health behavior change interventions and aligning the methods they chosen with relevant public health, work place or research goals.

INTRODUCTION
The main causes of morbidity and mortality (heart disease, cancer, etc.) are influenced by multiple health behaviors [1]. A majority of adults meet criteria for two or more risk behaviors [2, 3]. The Institute of Medicine estimated that 60,000 cancer deaths and 100,000 cases of cancer could be prevented in the USA each year by 2015, if more people adhered to the cancer prevention recommendations currently available [4]. These recommendations are also applicable to many common chronic diseases in the USA and globally.

Although disease prevention/health promotion interventions increasingly target more than one risk factor [5–8], there has been a relatively limited emphasis on the exploration of change across risk factors. Thus, few interventions that target multiple risk behaviors have examined those outcomes in combination [5, 9], instead examining change in individual risk factors. Increased focus on multiple outcomes could improve these interventions, for example, by determining which combinations of risk factors might yield the highest population-level impact on chronic disease prevention.

To date, reporting behavior change for each behavior separately has been the most common approach to reporting multiple behavior change intervention results; however, by focusing only on single risk factors, there is no ability to estimate the overall effect of the intervention. More studies are extending this method by calculating a simple *summative index* to reflect number of target behaviors, assigning a point for each behavior that meets public health recommendations [10]. The change in behaviors meeting recommendations pre- and post-intervention is assessed. Further, this outcome can be easily used to address Healthy People 2020 goals and other national recommendations; however, dichotomizing behaviors decreases sensitivity because it does not allow for continuous measurement. Thus, this approach does not allow for observations of incremental change or clustering of change in behavioral factors, both of which could be important for health outcomes and the assessment of the intervention's effectiveness [10].

However, there are other methods of assessing change in multiple behavior change interventions that have been suggested more recently [10, 11]. To create a *combined change score*, a z-score is calculated to standardize change in behaviors such that categorical and continuous variables can be combined into one score. However, each behavior is given equal weight, and the results may be limited in assessing meaningful changes in health behaviors [10]. Another approach is *optimal linear combination* that models a linear combination of multiple behavioral risk factors, including interactions between risk factors [11]. This method applies empirically driven weights to risk factors based on factors that are more likely to be changed by the intervention and allows for the examination of the effects of clustering; however, the use of the weights limits the ability of the results to be generalized to other populations [11]. The final approach is estimating the *impact of the intervention* based on the percent of those at risk and the percent of those who change that risk, which provides an indication of the impact of the intervention on both the participant and population levels by incorporating reach and efficacy [10, 12]. However, this method has not been used often by multiple health behavior change interventions; therefore, there is little guidance as to what constitutes a large impact.

The method(s) chosen to analyze multiple behavior change interventions may depend on intervention design, implementation, or study population. Given the descriptions above of multiple health behavior change methods, we have selected four methods to present here: (1) individual and summative behavior change, (2) combined change z-score, (3) optimal linear combination score, and (4) impact score. We utilized data from two cluster-randomized trials of large populations (health centers and small businesses). These trials targeted the same health behaviors in two different settings and provided the opportunity to examine how these methods work in interventions using two different channels. To our knowledge, this is the first paper to assess more than one method of analyzing multiple behavior change within a single dataset. Our goal is to provide insight on the comparability of different behavior change methods and to discuss the methodological and theoretical considerations in choosing methods for multiple health behavior change interventions.

METHODS

We analyzed baseline and follow-up data from two randomized controlled intervention trials in the Harvard Cancer Prevention Program Project: small business [8, 13] and health centers [7, 14]. These sister studies were conducted in parallel between 1999 and 2003 using similar intervention methods and simultaneously targeted four different health behaviors–fruit and vegetable consumption, red meat intake, leisure-time physical activity, and multivitamin use. Details of the study design [15, 16] and main outcome results [7, 8] have been published. A brief description is presented here.

Small business study

Investigators identified manufacturing worksites in Eastern Massachusetts using the Dun and Bradstreet database (Standard Industrial Classification codes 20–39). Worksite eligibility criteria included the following: employing a multicultural/multiethnic population, employing between 30 and 150 employees, having a turnover rate of less than 20 % in the previous year, and being autonomous in decision-making power to participate in a study (if part of a parent company). Of 224 manufacturing worksites, 197 (88 %) completed the pre-recruitment survey, 131 (66 %) met eligibility criteria, and 26 (12 %) consented to participate in the trial. Thirteen sites were randomized to the intervention condition, and 13, to the control condition. Two sites (one from each intervention condition) dropped out of the study, citing lack of time. Details of the recruitment process and a comparison of participating versus nonparticipating worksites can be found elsewhere [17].

The intervention strategies were based on (1) principles of employee participation and (2) a social contextual framework [18] that focused on multiple levels of influence on behaviors, with special attention given to work experiences, low literacy skills, and unique features of culture between ethnic groups, as well as the shared themes across cultural groups. Over the 18-month intervention period, monthly intervention activities focused on a change in the four targeted behaviors, and monthly contacts were made with management regarding environmental support and organizational change in each of the 12 intervention worksites [19].

Data were collected by interviewer-administered surveys of two (baseline and follow-up) cross-sectional samples of workers. The response rate in the 24 sites at final assessment was 77 % (range=54–93 %, n=1,408). Participation in the follow-up survey was not contingent on participation in the intervention. An embedded cohort of 974 participants in 24 worksites completed both baseline and final surveys. We present analyses of the embedded cohort.

Health centers study

Health centers were recruited for study participation from Harvard Vanguard Medical Associates, a health care system comprising 14 multispecialty community health centers in greater Boston. All ten health centers (100 %) invited to enroll agreed to participate. This was a cluster-randomized trial, with recruitment and randomization at the health center level. Providers practicing in the internal medicine department of each center were asked for permission to recruit from their patient pool; 83 % of providers (n=97) agreed. Participant eligibility criteria include the following: live in a low-income, multiracial/multiethnic neighborhood (based on geocoding); had a well or follow-up care visit scheduled with a participating provider; spoke and read either English or Spanish; had not been diagnosed with cancer; were not employed by any of the participating health centers or worksites; and consented to participate in the randomized control trial [7].

The intervention used the social contextual framework [18] and consisted of (1) provider endorsement of the intervention and a prescription for the recommended health behavior changes, (2) an in-person counseling session with a health advisor, (3) four follow-up telephone counseling sessions, (4) six sets of tailored materials that targeted social contextual factors, and (5) connections to relevant community activities. The intervention was offered in English and Spanish, and 90 % of intervention participants completed at least five of the six intervention activities.

Participants were enrolled by completing a telephone survey 1 week prior to health care visit. Study staff attempted to recruit 8,963 potentially eligible candidates. Of these, 2,547 (28 %) were unreachable; 867 (10 %) were ineligible; 3,330 (37 %) refused; and 2,219 (25 % of those reached; 40 % of those eligible) were enrolled [7]. All those who completed the baseline survey were contacted for a follow-up survey, of which the response rate was 88 % and was similar across conditions.

MEASURES

Health behaviors

Servings of fruits and vegetables (FV) consumed each day were assessed with a screener developed for the National Cancer Institute's nine 5-A-Day for Better Health research studies [20]. Responses were re-coded to equivalent servings and summed to obtain total servings of fruits and vegetables per day, expressed as a continuous variable.

Red meat consumption (RM) was assessed with an abbreviated form of a semiquantitative food frequency questionnaire [21]. The responses were recoded to equivalent servings per week and summed for total servings of red meat per week, expressed as a continuous variable.

The leisure-time physical activity (PA) assessment was based on the questionnaire used in the Nurses' Health Study [22], adapting items to specific activities that were found to be more common among the intended population (e.g., omitting tennis and adding dance). Respondents were asked to indicate how often, on average, over the past 4 weeks they had engaged in eight moderate- or vigorous-level physical activities. In addition, respondents were asked about usual walking pace. The responses were re-coded from METs to equivalent minutes per week and summed to yield total minutes of physical activity per week. Walking was included if usual walking pace was reported to be faster than "easy, casual."

Multivitamin use (MV) was assessed with a single question asking workers how many days per week, on average, they took a multivitamin. The variable was coded 0–7 with 0 indicating the participant never took multivitamins and 1–7 indicating average number of days per week the participant took a multivitamin. Although there have been mixed results, studies have shown multivitamin use as protective for cancer and other chronic diseases [23, 24].

ANALYSIS

Method 1: summative index

We dichotomized each health behavior variable according to national recommendations and report the percentage of respondents who changed from not meeting the intervention recommendation at baseline to meeting the recommendation at follow-up. Health behaviors were dichotomized as follows: for FV consumption, 1= meeting recommendations (e.g., consuming five or more servings of fruits and vegetables per day) and 0= not meeting recommendations (e.g., less than five per day) [25]. For RM intake, 1=meeting recommendations (e.g., consuming three or fewer servings of red meat per week) and 0=not meeting recommendations (consuming more than three per week) [25–31]. For PA, 1= meeting recommendations (participating in two and a half hours or more per week of leisure-time physical activity) and 0=not meeting recommendations (participating in less than two and a half hours per week of PA) [26, 32]. For MV, 1=meeting recommendation (multivitamin use at least 6 days per week) and 0=not meeting recommendation (multivitamin use less than 6 days per week) [23, 24]. We adjusted for clustering of respondents by site via multilevel logistic regression, modeling site at level 2 and individuals at level 1. In these equations, each health behavior was an outcome in separate models; intervention status was the predictor. Change in each behavior was summed for each participant, and the average percent change was calculated. The χ^2 test at alpha=0.05 was used to compare the intervention to the control group.

Method 2: combined change z-score

We summed behavior-specific standardized change scores into an index of overall behavior change. To create standardized change scores for each behavior, we subtracted baseline values from follow-up values of continuously coded behavior variables and divided by the standard deviation of the difference. To create a behavior change index, we added the resulting z-scores for FV consumption, PA, and MV use and then subtracted the z-score for RM consumption, since the intervention sought to decrease this behavior. To test for significant change between the intervention and control group, we used a two-tailed t test with an alpha=0.05.

Method 3: optimal linear combination

In this approach, proposed by Goodman and colleagues [11], we calculated a summary behavior change score from an optimal score function for the four standardized health behaviors. To keep all variables on an equivalent time scale (per week), we created a new variable for FV consumption that calculated the amount of fruits and vegetables consumed in 1 week by multiplying the current measure by seven. Red meat consumption was reverse coded so that its recommended direction of change was positive. Next, we standardized each continuously coded health behavior using the following formula:

$ST(B) = (B - P_{05})/(P_{95} - P_{05})$ where B is the behavior's original value, P_{05} and P_{95} are the 5th and 95th percentile values, respectively, for B, and ST (B) is the new standardized behavior variable: ST (FV), ST (RM), ST (PA), and ST (MV). We used multilevel logistic regression to find an optimal linear combination to assess how the standardized behavioral variables would predict intervention status. In these models, the standardized behavior variables at baseline and follow-up are the predictors, and observations are nested within individuals. We used the resulting significant β-coefficients for the health behaviors at follow-up to create a multiple behavior change summary score (main effects model) for the small business study: Index Score A=0.98×ST (FV)+0.77×ST (PA)+0.57×ST (MV) and the health centers study: Index Score A=1.08×ST (FV)+0.72×ST (RM)+1.56×ST (MV).

In the final score function, we multiply the effects (parameter estimates) by 100 to increase the range of the scores as well as to simplify interpretation. Significance was tested using two-sided t tests assuming equal variances at alpha=0.05.

Method 4: expanded intervention impact score

In this approach, we quantify multiple behavior change with an expanded intervention impact formula. Impact traditionally has been measured for interventions that target a single behavior as efficacy multiplied by participation ($I=E\times P$) [12, 33] but can be measured for interventions that target multiple behaviors as efficacy multiplied by participation summed over all targeted behaviors [$I=\sum_{\text{no. of behavior }(n)} (E_n\times P_n)$] [12, 34]. P is the proportion of individuals participating in the intervention who are at risk for each behavior (the percentage of the study population not meeting recommended levels of target behaviors at baseline). E is the estimate of efficacy for each behavior (the percentage of the study population meeting recommended levels of target behaviors at follow-up). Because this method requires a measure of participation, it is not measured in control groups and, thus, typically assessed among intervention participants only. However, if there was a trial with two active intervention conditions, comparisons between the two could be made using the χ^2 test and α=0.05 to compare the two intervention groups. As described by Prochaska and colleagues [10], impact values for single behaviors range from 0 to 1, but impact values for multiple behaviors can exceed 1. Therefore, interventions targeting more behaviors have the potential for greater impact. There is no significance testing used for this method.

RESULTS

Tables 1 and 2 present the four methods of behavior change results for the small business and health center studies, respectively. The results of method 1 have been published elsewhere [7, 8] and are shown here for comparison with the other three methods.

Small business study

Table 1 presents all four methods of health behavior change results for the small business study. The results of method 1 (summative) revealed an overall significant summative index (p<0.0001). The average change per individual for all health behaviors was significantly greater in the intervention group compared to the control group. Using method 2 (z-score), respondents in the intervention group had larger z-scores than those in the control group for the index of multiple behavior change and for each behavior, indicating more improvement. Method 2 was the only method to show significant change in PA (p=0.03). The difference in improvement was statistically significant for z-score index (p<0.0001), FV consumption (p=0.03), PA (p=0.03), and MV use (p<0.0001), but not for RM (p=0.81). The difference between baseline and follow-up for the combined behavior change index for method 3 (linear combination) was positive among intervention group respondents, indicating change in the recommended direction. Compared to the control group, behavior change in the intervention group was statistically significant for FV consumption (p<0.0001), MV use (p<0.0001), and the combined behavior change index (p<0.0001). The largest individual impact of the intervention, method 4, was on MV intake (0.27); however, the impact of the other health behaviors was only slightly lower. The total population impact is 0.94.

Health centers study

Table 2 presents all four methods of behavior change results for the health centers study. The results of method 1 (summative) revealed an overall significant change in the intervention group compared to the control group (p<0.0001). Similar results were seen from methods 2 (z-score) and 3 (linear combination). Results from method 2 (z-score) showed that respondents in the intervention group had larger z-scores and a statistically significant improvement compared to those in the control group for the index of multiple behavior change (p<0.0001), FV consumption (p<0.0001), RM (p<0.0001), and MV use (p<0.0001), but not for PA (p=0.77). Results from method 3 (linear combination) showed that the intervention group had positive values for the difference in standardized behavior scores between baseline and follow-up, indicating change in the recommended direction. The most prevalent risk behavior, shown in method 4 (impact), was low FV consumption, and the least prevalent risk behavior was low PA. As demonstrated by the "individual impact" scores, the intervention had the most impact on MV use (0.42), followed by RM (0.31), and the least impact on FV consumption (0.15). The total population impact is 1.08.

DISCUSSION

Using four evaluation methods, we explored the effectiveness on multiple health behavior changes of

Table 1 | Results of methods of behavior change, small business study

	Group		
Method	Control	IX	*p* value
Method 1: summative index	% change	% change	
Fruit and vegetables	0.80	7.60	<0.0001
Red meat	4.70	5.40	0.41
Leisure-time physical activity	−2.10	6.40	0.25
Multivitamin intake	0.80	10.10	0.01
Combined summative index	−0.03	−0.28	<0.0001
Method 2: combined change z-score	Mean (SE)	Mean (SE)	
Fruit and vegetables	0.03 (0.05)	0.19 (0.05)	0.03
Red meat	−0.15 (0.05)	−0.17 (0.06)	0.81
Leisure-time physical activity	−0.12 (0.08)	0.13 (0.08)	0.03
Multivitamin intake	0.04 (0.04)	0.26 (0.05)	<0.0001
z-Score index	0.04 (0.12)	0.82 (0.13)	<0.0001
Method 3: optimal linear combination	Score	Score	
Fruit and vegetables	−3.25	0.75	<0.0001
Red meat	−1.80	−2.00	0.38
Leisure-time physical activity	−4.36	0.97	0.11
Multivitamin intake	1.06	6.53	<0.0001
Combined index	−5.57	7.86	<0.0001
Method 4: impact score	$P \times E$	Impact	
Proportion at risk × Efficacy = Impact			
Fruit and vegetables	(85.18) (22.37)	0.19	
Red meat	(69.09) (35.46)	0.24	
Leisure-time physical activity	(29.33) (77.43)	0.23	
Multivitamin intake	(69.67) (39.34)	0.27	
Population impact			0.94

P proportion at risk, *E* efficacy, *SE* standard error

two cluster-randomized trials of large multiracial/multiethnic populations conducted in parallel, in two different settings, the worksite and health centers. Method 1 (summative) is useful for examining the ability of the intervention to encourage participants to meet public health recommendations and gives equal value to each behavior targeted. Methods 2 (z-score) and 3 (linear combination) allow us to look at a change in each behavior on a continuous scale in order to assess incremental change. Method 4 (impact) examines intervention impact for individual and multiple behaviors and allows for the comparison of intervention impact across studies with different designs and target behaviors. As described in Table 3, methods 2, 3, and 4 allow for comparison of the effectiveness in changing health behaviors between more than one intervention approaches. The method used to calculate a change in the individual behaviors influences the overall behavior change score. Based on significance testing, each method produced similar results for overall health behavior change scores. Methods 1, 2, and 3 showed a magnitude increase in physical activity in the intervention group from baseline to follow-up and a reduction in the control group. Method 4 is only calculated among the intervention group. Method 2 (z-score) was able to detect significantly more change in the intervention group compared to the control group, whereas the other methods did not. Both methods 2 (z-score) and 3 (linear combination) were able to assess continuous, incremental increase in physical activity; however, method 3 (linear combination) applied weights to the health behaviors and controlled for clustering of behaviors which reduced the significance of the change in physical activity among the intervention group.

Methodological considerations

There are a number of methodological decisions to consider when designing and planning the analysis of multiple health behavior change interventions. When assessing continuous health behavior change variables, standardization techniques may become necessary to minimize the influence of outliers and to make a one unit change in one variable similar to a one unit change in another. Additionally, while both methods 2 (z-score) and 3 (linear combination) construct linear combinations of individual components to summarize the overall effect of the intervention, they differ by the weights used in their linear constructs. Method 2 (z-score) considers an unweighted linear combination, placing equal weight on each combination, whereas the weights used in method 3 (linear combination) are empirically driven, specifically targeting the components that are more likely to be changed by the intervention.

Assessing behavior change with methods 1–3 allows us to assess the significance of the change

Table 2 | Results of methods of behavior change, health centers study

Method	Group Control	IX	*p* value
Method 1: summative index	% change	% change	
Fruit and vegetables	−4.00	3.60	<0.0001
Red meat	0.40	11.20	<0.0001
Leisure-time physical activity	0.80	1.20	0.6
Multivitamin intake	7.50	29.50	<0.0001
Combined summative index	−0.08	−0.31	<0.0001
Method 2: combined change z-score	Mean (SE)	Mean (SE)	
Fruit and vegetables	−0.04 (0.03)	0.18 (0.03)	<0.0001
Red meat	−0.05 (0.03)	−0.26 (0.03)	<0.0001
Leisure-time physical activity	−0.02 (0.02)	−0.02 (0.02)	0.81
Multivitamin intake	0.18 (0.03)	0.63 (0.03)	<0.0001
z-Score index	0.2 (0.06)	1.05 (0.06)	<0.0001
Method 3: optimal linear combination	Score	Score	
Fruit and vegetables	−2.93	4.51	<0.0001
Red meat	−2.03	3.45	<0.0001
Leisure-time physical activity	0.03	0.06	0.77
Multivitamin intake	13.35	47.12	<0.0001
Combined index	8.54	55.08	<0.0001
Method 4: impact score	$P \times E$	Impact	
Proportion at risk × Efficacy = Impact			
Fruit and vegetables	(86.36) (17.00)	0.15	
Red meat	(51.66) (60.00)	0.31	
Leisure-time physical activity	(29.50) (71.86)	0.21	
Multivitamin intake	(60.52) (68.75)	0.42	
Population impact			1.08

P proportion at risk, *E* efficacy, *SE* standard error

between baseline and follow-up by comparing change in the intervention group with change in the control group. However, method 4 (impact) specifies calculating impact for the intervention population only, resulting in a lack of comparison population. Impact equations can include all behaviors that are changed, including behaviors targeted directly through high- or low-intensive interventions or behaviors that changed even if they were not directly targeted by intervention activities [10]. Additionally, low prevalent behaviors contribute less to the impact score, even if they are major risk

Table 3 | Sample behavioral intervention research questions assessed via multiple evaluation methods

Research question	Methods 1	2	3	4
Intervention types				
1. Is MHBC intervention approach A more effective than MHBC intervention approach B in changing an individual behavior?	X			
2. Is MHBC intervention approach A more effective than MHBC intervention approach B in changing all the targeted behaviors?		X	X	X
3. Is MHBC intervention approach A more effective than MHBC intervention approach B in changing all the targeted behaviors, accounting for interactions between variables?			X	
Behavioral targets				
4. Did the delivery of this MHBC program result in attainment of health goals (e.g., Healthy People 2010 goals) for any of the targeted behaviors?	X			X
5. Did the delivery of this MHBC program result in incremental behavioral change for any of the targeted behaviors?		X	X	
6. What would the expected effect on disease prevention for my organization as a whole be if we were to implement this MHBC program?				X

Methods 1 (summative index), 2 (combined change z-score), 3 (optimal linear combination), and 4 (impact) were used *MHBC* multiple health behavior change factors, such as tobacco use [35]. As more researchers assess the population impact score of their interventions, it will become necessary to carefully evaluate the meaning of these scores for two reasons: the population impact score of all multiple behaviors can exceed 1, and there is no upper limit for this score. The results of these future studies will begin to increase the ability to compare one's impact score between interventions; however, more research is needed on interpreting population impact and assessing its significance.

Practical considerations

Our results provide practical considerations to researchers to guide their decision in the selection of a multiple behavior change evaluation method. In general, if an innovative approach to multiple behavior change is being tested or if a clear consistent effort was made to integrate the behaviors within the intervention activities, methods 2 (z-score), 3 (linear combination), and 4 (impact) should be reported to specifically report efficacy in multiple behavior change. However, if presentation of behaviors was administered in a sequential format with little attention paid to integrating behaviors, reporting change in individual levels of behaviors may be sufficient. Drawing from the recognized need to demonstrate outcomes that are important for decision making, we believe that the selection of an evaluation technique should be suited to the intended audience and setting [36]. In Table 3, we illustrate how different research questions may be suited to the different evaluation methods. For example, methods 1 (summative), 2 (z-score), and 3 (linear combination) are capable of comparing an intervention condition to a control condition, thereby establishing efficacy. As another example, aligning results with relevant public health goals is an important consideration for translating interventions to practice settings; thus, methods 1 (summative) and 4 (impact) would be well suited. However, in other cases, other goals may be considered to be of higher priority than establishing compliance with public health recommendations. For example, workplace research may value increases in employee productivity and decreases in health care costs as much as beneficial behavioral changes. In this case, demonstrating incremental progress toward achieving these goals of behavioral change using methods 2 (z-score) and 3 (linear combination) may be what is most valuable.

Because the interventions tested here included continuous behavior change variables, we were able to analyze all four methods and compare the interpretations of each. However, when interventions include behaviors that are dichotomous (e.g., smoking cessation or receipt of screening), our findings would suggest that methods 1 (summative) and 4 (impact) would be appropriate techniques for evaluation. Future directions for this type of research may include application of these evaluation methods to not only interventions focused on behavior change but also the effects of change on disease outcomes. For example, King and colleagues [37] examined the adoption of a healthy lifestyle with mortality and cardiovascular disease using a summative index of four lifestyle variables. Future research may expand to determine associations using the additional evaluation techniques reported here.

CONCLUSIONS

Researchers should move beyond individual outcome reporting in multiple health behavior change interventions for at least three reasons: (1) principles underlying change in some behavioral risk factors for chronic disease are similar, and thus, our interventions might have more impact than it appears based on individual risk factor outcomes; (2) multiple significance testing may inflate type I error [10]; and (3) understanding the impact of interventions across risk factors could increase research dissemination and implementation. The more researcher will use and report appropriate methods for their multiple health behavior change interventions, the closer the field of multiple behavior change will move toward a consensus on the use of these methods and markers of success.

Acknowledgments: This work was conducted with the support of a grant from the National Cancer Institute (grant P01 CA 75308). Dr. Drake was supported by funding by the Barnes-Jewish Hospital Foundation.

1. Kilmer G, et al. Surveillance of certain health behaviors and conditions among states and selected local areas—Behavioral Risk Factor Surveillance System (BRFSS), United States, 2006. *MMWR Surveill Summ*. 2008;57(7):1-188.
2. Fine LJ, et al. Prevalence of multiple chronic disease risk factors. 2001 National Health Interview Survey. *Am J Prev Med*. 2004;27 (2):18-24.
3. Pronk NP, et al. Meeting recommendations for multiple healthy lifestyle factors. Prevalence, clustering, and predictors among adolescent, adult, and senior health plan members. *Am J Prev Med*. 2004;27(2):25-33.
4. Curry SJ, Byers T, Hewitt M, eds. *Fulfilling the potential for cancer prevention and early detection*. National Academy Press; 2003: 542
5. Emmons K, et al. Project PREVENT: a randomized trial to reduce multiple behavioral risk factors for colon cancer. *Cancer Epidemiol Biomarkers Prev*. 2005;14(6):1453-9.
6. Emmons KM, et al. Prevalence and predictors of multiple behavioral risk factors for colon cancer. *Prev Med*. 2005;40 (5):527-34.
7. Emmons KM, et al. Cancer prevention among working class, multi-ethnic adults: results of the healthy directions health centers study. *AJPH*. 2005;95(7):1200-5.
8. Sorensen G, et al. Promoting behavior change among working-class, multi-ethnic workers: results of the healthy directions small business study. *American Journal of Public Health*. 2005;95(8):1389-1395.
9. King TK, et al. Cognitive-behavioral mediators of changing multiple behaviors: smoking and a sedentary lifestyle. *Prev Med*. 1996;25(6):684-91.
10. Prochaska JJ, et al. Methods of quantifying change in multiple risk factor interventions. *Prev Med*. 2008;46(3):260-5.
11. Goodman MS, et al. An evaluation of multiple behavioral risk factors for cancer in a working class, multi-ethnic population. *Journal of Data Science*. 2006;4:291-306.

12. Abrams DB, et al. Integrating individual and public health perspectives for treatment of tobacco dependence under managed health care: a combined stepped-care and matching model. *Ann Behav Med*. 1996;18(4):290-304.
13. Hunt MK, et al. Cancer prevention for working class, multiethnic populations through small businesses: the healthy directions study. *Cancer Causes Control*. 2003;14(8):749-60.
14. Emmons KM, et al. Cancer prevention for working class, multiethnic populations through health centers: the healthy directions study. *Cancer Causes Control*. 2003;14(8):727-37.
15. Kim YI. Folate and DNA methylation: a mechanistic link between folate deficiency and colorectal cancer? *Cancer Epidemiol Biomarkers Prev*. 2004;13(4):511-9.
16. Giovannucci E. Epidemiologic studies of folate and colorectal neoplasia: a review. *J Nutr*. 2002;132(8 Suppl):2350S-2355S.
17. Barbeau EM, et al. Recruiting small manufacturing worksites that employ multiethnic, low-wage workforces into a cancer prevention research trial. *Prev Chronic Dis*. 2004;1(3):A04.
18. Sorensen G, et al. Model for incorporating the social context in health behavior intervention: applications for cancer prevention for working-class, multi-ethnic populations. *Preventive Medicine*. 2003;37(3):188-97.
19. Rosenstock I, ed. *The health belief model: explaning health behavior through expectancies. Health behavior and health education: theory, research and practice*. San Francisco, CA: Jossey-Bass; 1990.
20. Subar AF, et al. Fruit and vegetable intake in the United States: the baseline survey of the Five A Day for Better Health Program. *Am J Health Promot*. 1995;9(5):352-60.
21. Schatzkin A, et al. Diet and colorectal cancer: still an open question. *J Natl Cancer Inst*. 1995;87(23):1733-5.
22. Wolf AM, et al. Reproducibility and validity of a self-administered physical activity questionnaire. *Int J Epidemiol*. 1994;23(5):991-9.
23. Giovannucci E, et al. Multivitamin use, folate, and colon cancer in women in the nurses' health study. *Ann Intern Med*. 1998;129 (7):517-24.
24. Jacobs EJ, et al. Multivitamin use and colon cancer mortality in the cancer prevention study II cohort (United States). *Cancer Causes Control*. 2001;12(10):927-34.
25. World Cancer Research Fund. *Food, nutrition, physical activity, and the prevention of cancer: a global perspective*. Washington, DC: AICR; 2007.
26. Surgeon General's report on physical activity and health. From the Centers for Disease Control and Prevention. *Jama*. 1996: 276 (7): 522
27. Guidelines on diet, nutrition, and cancer prevention: reducing the risk of cancer with healthy food choices and physical activity. The American Cancer Society 1996 Advisory Committee on Diet, Nutrition, and Cancer Prevention. *CA Cancer J Clin*. 1996; 46(6): 325–41
28. de Verdier G. M., et al., Meat, cooking methods and colorectal cancer: a case-referent study in Stockholm. *Int J Cancer*. 1991;49 (4):520-5.
29. Giovannucci E, et al. Intake of fat, meat, and fiber in relation to risk of colon cancer in men. *Cancer Res*. 1994;54(9):2390-7.
30. Sandhu MS, White IR, McPherson K. Systematic review of the prospective cohort studies on meat consumption and colorectal cancer risk: a meta-analytical approach. *Cancer Epidemiol Biomarkers Prev*. 2001;10(5):439-46.
31. Willett WC, et al. Relation of meat, fat, and fiber intake to the risk of colon cancer in a prospective study among women. *N Engl J Med*. 1990;323(24):1664-72.
32. Jacques PF, et al. The effect of folic acid fortification on plasma folate and total homocysteine concentrations. *N Engl J Med*. 1999;340(19):1449-54.
33. Velicer WF, Prochaska JO. An expert system intervention for smoking cessation. *Patient Educ Couns*. 1999;36(2):119-29.
34. Prochaska JJ, et al. Comparing intervention outcomes in smokers treated for single versus multiple behavioral risks. *Health Psychol*. 2006;25(3):380-8.
35. Shafey O, Dolwick S, Guindon GE. *Tobacco control country profiles 2003*. Atlanta, GA: American Cancer Society; 2003.
36. Chang MH, et al. Decreased incidence of hepatocellular carcinoma in hepatitis B vaccinees: a 20-year follow-up study. *J Natl Cancer Inst*. 2009;101(19):1348-55.
37. King DE, Mainous AG, Geesey ME. Turning back the clock: adopting a healthy lifestyle in middle age. *The American Journal of Medicine*. 2007;120:598-603.