

Published in final edited form as:

Am J Prev Med. 2007 May ; 32(5 Suppl): S112–S118.

The Multiphase Optimization Strategy (MOST) and the Sequential Multiple Assignment Randomized Trial (SMART): New Methods for More Potent eHealth Interventions

Linda M. Collins, PhD¹, Susan A. Murphy, PhD², and Victor Strecher, PhD³

1 *The Methodology Center and Department of Human Development and Family Studies, The Pennsylvania State University, University Park, PA*

2 *Institute for Social Research and Department of Statistics, University of Michigan, Ann Arbor, MI*

3 *Department of Health Behavior and Health Education, University of Michigan, Ann Arbor, MI*

Abstract

In this article two new methods for building and evaluating e-health interventions are described. The first is the Multiphase Optimization Strategy (MOST). MOST consists of a screening phase, in which intervention components are efficiently identified for selection for inclusion in an intervention or rejection, based on their performance; a refining phase, in which the selected components are fine-tuned, and questions such as optimal component dosage are investigated; and a confirming phase, in which the optimized intervention, consisting of optimal doses of the selected components, is evaluated in a standard randomized confirmatory trial. The second is the Sequential Multiple Assignment Randomized Trial (SMART) which is an innovative research design especially suited for building time-varying adaptive interventions. A SMART trial can be used to identify the best tailoring variables and decision rules for an adaptive intervention empirically. Both the MOST and SMART approaches use randomized experimentation to enable valid inferences. When properly implemented, these approaches will lead to the development of more potent e-health interventions.

There are good reasons to believe that interventions based on e-health principles have the potential for considerable public health impact. Perhaps the most obvious reason is the reach of these interventions. Once an electronic intervention has been designed and programmed, delivery occurs via methods such as the Internet or by mailing a CD, and therefore is extremely convenient. Moreover, the incremental cost of delivering an intervention to additional people is usually negligible, certainly in comparison to traditional interventions where in order to reach more recipients it becomes necessary to add additional physicians, therapists, health educators, peer counselors, and so on to deliver the program. The limiting factor for reach of an e-intervention is less likely to be a shortage of resources for delivering the program electronically than access to computers on the part of potential recipients. However, access to computers continues to increase in all strata of American society, suggesting that e-health interventions hold growing promise.¹

Correspondence and reprint requests: Linda M. Collins The Methodology Center, Penn State 204 East Calder Way, Suite 400, State College, PA 16801 Phone: (814) 865–7090; Fax: (814) 863–0000; E-mail: LMCollins@psu.edu

Publisher's Disclaimer: This is a PDF file of an unedited manuscript that has been accepted for publication. As a service to our customers we are providing this early version of the manuscript. The manuscript will undergo copyediting, typesetting, and review of the resulting proof before it is published in its final citable form. Please note that during the production process errors may be discovered which could affect the content, and all legal disclaimers that apply to the journal pertain.

No financial conflict of interest was reported by the authors of this paper.

The broad reach of e-health interventions is particularly exciting in the light of some new methods for building and optimizing behavioral interventions. The purpose of this article is to introduce two of these new methods to e-health scientists. One is the Multiphase Optimization Strategy (MOST)² for building and evaluating interventions in such a way that they are made out of active program components delivered at optimal doses. The other is the Sequential Multiple Assignment Randomized Trial (SMART)³ for building adaptive interventions. We propose that these new methods, although relatively untried at this writing, are eminently practical and hold much potential for e-health research. By using these methods, it is possible to produce more potent interventions which, when coupled with the reach afforded by e-health approaches, will promise considerable overall public health impact.

THE MULTIPHASE OPTIMIZATION STRATEGY (MOST)

The traditional approach to intervention development has involved constructing an intervention a priori and then evaluating it in a standard randomized confirmatory trial (RCT). After the confirmatory trial, post-hoc analyses are done to help explain how the intervention worked, or why it did not work. The results of these analyses may be used to refine the intervention program and construct a second generation version of the program, which is then evaluated in a new RCT.

Collins, Murphy, Nair, and Strecher² reviewed shortcomings of this approach. While acknowledging that RCTs are the undisputed gold standard for assessing the effect of an intervention as a package once it has been developed, they pointed out that the post hoc analyses that typically follow a RCT in order to inform further intervention design and evaluation are subject to bias because they are not based on random assignment. As a result the cycle of intervention – RCT – post hoc analyses – revision of intervention – RCT is likely to lead very slowly, if at all, to an optimized intervention.

Collins et al. also pointed out that most behavioral interventions can be considered an aggregation of a set of components. Some intervention components are a part of the program itself (e.g. program content). Others may be more concerned with the delivery of the program (e.g., whether a message is delivered by a lay person or by a physician). Some components may be having the intended effect; others may be having no effect at all; and others may even be reducing the overall potency of the intervention. Because the traditional RCT evaluates the intervention only as a whole, using the RCT alone does not enable isolation of the effects of individual program or delivery components. A different experimental approach is necessary to accomplish this.

RCT: randomized confirmatory trial

We suggest MOST as an alternative way of building, optimizing and evaluating e-health interventions. MOST incorporates the standard RCT, but *before the RCT is undertaken* also includes a principled method for identifying which components are active in an intervention, and which doses of each component lead to the best outcomes. The principles underlying MOST are drawn from engineering, and emphasize efficiency. MOST consists of three phases, each of which addresses a different set of questions about the intervention *by means of randomized experimentation*.

Figure 1 offers an outline of the three phases of MOST. The first phase is *screening*. The starting point for the screening phase is a previously identified finite set of intervention components, made up of program components and/or delivery components. It is assumed that there is some theoretical basis for the choice of these components. It is also assumed that any initial pilot testing necessary to assess feasibility and finalize the details of implementation has been completed prior to the start of the screening phase.



**START; PHASE 1: Online - Fragebogen / PHASE 2: Ernährung / PHASE 3: Sport / PHASE 4: Lifestyle / PHASE 5: Feedback
Theoretischer Hintergrund:**

The objective of the screening phase is to address questions like the following: Which of the set of program components are active and contributing to positive outcomes, and should be included in the intervention? Which program components are inactive or counterproductive, and should be discarded? Which of the set of delivery components are active and make a difference in the intervention outcome, and thus play a role in maintaining intervention fidelity? Decisions about which program and delivery components are active and should be retained and which are inactive and should be discarded are made based on the results of a randomized experiment. (Experimental design alternatives are discussed below.) The decision may be made on the basis of statistical significance at any alpha level deemed appropriate, or on the basis of estimated effect size. In addition, cost in relation to incremental contribution to the desired outcome may be a consideration. At the conclusion of the screening phase, a set of program and delivery components that are to be retained for further examination has been identified. This set of components constitutes a “first draft” intervention.

This “first draft” intervention is the starting point for the next phase of MOST, the *refining* phase. In this phase the “first draft” intervention is examined further, with the objective of fine-tuning the intervention and arriving at a “final draft.” The specific activities of the refining phase depend on the intervention being considered, but in general focus on questions such as: Given the components identified in the screening phase, what are the optimal doses? Does the optimal dose vary depending on individual or group characteristics? As in the screening phase, in the refining phase decisions are based on randomized experimentation, and cost may be a consideration. At the conclusion of the refining phase, the investigator has identified an optimized “final draft” intervention consisting of a set of active program and delivery components at the best doses.

The “final draft” intervention provides the starting point for the third phase of MOST, *confirming*. In the confirming phase this optimized intervention is evaluated in a standard RCT. The confirming phase addresses questions such as: Is this intervention, as a package, efficacious? Is the intervention effect large enough to justify investment in community implementation?

A brief illustrative example

Because MOST is an approach or a perspective rather than an off-the-shelf procedure, exact details about its implementation depend on the application. In order to illustrate MOST, we offer a brief hypothetical example similar to the one in Collins et al.². The example is based on (but not identical to or an account of) the work of one of the authors of the current article (VS).

Suppose the objective is to use MOST to build, optimize and evaluate an e-intervention for smoking cessation, and that six components have been identified for study, four of which are program components and two of which are delivery components. The program components are outcome expectation messages (messages addressing an individual’s expectations about what will happen if he or she quits smoking) which may be either present or absent in the intervention; efficacy expectation messages (these address barriers to perceived self-efficacy) which may be present or absent; message framing (this concerns how the persuasive messages about quitting smoking are to be framed), which may be positive or negative; and testimonials (from former smokers), which may be present or absent. The delivery components are exposure schedule, which may be one long message or four smaller ones; and source of message, which may be a primary care physician or the individual’s health maintenance organization (HMO).

In the screening phase randomized experimentation is conducted to isolate the effects of each of the six components. Suppose the experimental results indicate that the active program components are outcome expectation messages, efficacy expectation messages, and

testimonials, and that there is one active delivery component, exposure schedule. Once this “first draft” intervention has been identified, the screening phase is concluded. The intervention scientist now proceeds to the refining phase in order to fine-tune the “first draft” and arrive at an optimized intervention. An example of this fine-tuning might be experimentation to pinpoint the best dose of exposure schedule, in other words, the optimal number of messages. The “final draft” intervention would then consist of outcome expectation messages, efficacy expectation messages, and testimonials, with the intervention delivered using the optimal number of messages identified in the refining phase. In the confirming phase, this “final draft” smoking cessation intervention is evaluated in a standard RCT.

Design for the screening and refining phases

The research design to be used in the confirming phase (i.e., the RCT) is straightforward and familiar to most intervention scientists. Usually a simple two-group design consisting of random assignment to either a program condition or a suitable comparison condition would be used. It may be less evident what design would be used in the screening and refining phases. One family of designs that lends itself well to the screening and refining phases is the factorial analysis of variance (ANOVA). In an ANOVA design several independent variables, or factors, are investigated at once. A properly chosen and implemented ANOVA design permits the effects of individual independent variables to be isolated. In the behavioral sciences the factors are usually “fully crossed” which means that each level of a variable is combined with each level of the other variables.

For example, suppose there are just two program components under consideration: outcome expectation messages and efficacy expectation messages. To examine these in the screening phase using a fully crossed factorial ANOVA, subjects would be randomly assigned to one of four experimental conditions: both messages present; outcome expectation messages only; efficacy expectation messages only; and both messages absent (perhaps an information-only control). At the end of the screening phase, after the experiment was completed, the decision about which components to select for further consideration would be based on the main effect and interaction estimates obtained from the ANOVA. The decision may be made by selecting statistically significant effects; it may be made by choosing components associated with an estimated effect size over some threshold level; or it may use the results of the ANOVA in some other way.

Although factorial designs are the most efficient way to assess the effect of several independent variables simultaneously, they have for the most part been eschewed by intervention scientists because of the perception that they are impractical due to the number of conditions that must be implemented. For example, a fully crossed ANOVA design to investigate the six components in our example would involve 64 treatment conditions. This may in fact be too many conditions to manage for interventions delivered by teachers and practitioners in settings like schools and hospitals, but it does not necessarily follow that the field of e-health should be similarly discouraged about factorial ANOVA designs. Because e-health interventions are delivered electronically, the primary cost often will be the computer programming required to construct each of the conditions. Once this task is done, it may be relatively straightforward to assign individuals randomly to experimental conditions and then deliver the corresponding version of the intervention. Thus, factorial ANOVA may be more feasible in e-health than it is in other more traditional areas of intervention science.

However, when there are many factors, the construction of each condition in a fully crossed ANOVA design may be too much for an e-intervention study. In this case, fractional factorial ANOVA designs can be an attractive alternative. When using fractional factorial ANOVA designs, it is not necessary to include every possible experimental condition in the design. Instead, based on working assumptions made by the investigator, a subset of conditions is

chosen strategically in order to estimate effects of primary interest. Fractional factorial designs are not new; they go back to Fisher⁴ and Box, Hunter, and Hunter,⁵ and have been used routinely in engineering and agriculture for many years. Fagerlin et al.⁶ recently employed a fractional factorial approach in medical decision making research. Intervention science also can and should benefit from the efficiency and economy these designs provide.

Collins et al.² illustrated how a six-factor fully crossed ANOVA design with 64 conditions can be reduced to a fractional factorial ANOVA design with 16 conditions. The reduced design retains the capability to provide main effects estimates for each of the six independent variables, and also the capability to provide estimates for selected interactions. The power associated with the test of each main effect is the same as that for any simple two-group comparison. In the refining phase, variants on fractional factorial designs, such as response surfaces, may be useful for questions involving identification of optimal doses.

As mentioned above, in some situations it may not be necessary or desirable to base decisions strictly on hypothesis tests.² If hypothesis testing is to be used, it may be necessary to control the experiment-wise error rate. As a simple expedient, we suggest identifying a priori a limited set of effects predicted to be sizeable, testing those at the desired alpha level without regard to the experiment-wise error rate, and then using a Bonferroni or similar adjustment for the remaining effects. (For more about the experiment-wise error rate see Wu and Hamada.⁷) Note that in general interaction effect sizes tend to be small, making it important to power a study accordingly if interactions are of particular interest.

Although we propose that MOST is useful in a wide variety of intervention development settings, there are some situations in which investigators may wish to consider a different approach. When applied to the building of new interventions, MOST is based on the idea that it is feasible to identify individual program components that can stand alone, at least enough to assess their individual effects. It may not be sensible to parse an extremely tightly integrated program into separate parts. Even when meaningful individual program components can be identified, it may be expected that each component has a very small, difficult to detect effect that nevertheless contributes to a larger, more readily detectable cumulative effect of the entire package. If in addition it can safely be assumed that none of the components has a deleterious effect or reduces the efficacy of other components, then the effects of individual components may not be of much interest. However, MOST may still be helpful in examining delivery components associated with these interventions.

Even when an intervention can be meaningfully decomposed, it is possible that the list of components cannot be combined at will, in other words, not every combination of program components is sensible to implement. In some cases a fractional factorial design may be chosen that includes only sensible combinations of components. If there is a component that is expected not to operate properly in the absence of another component, it may be more fruitful to consider the two components as one for the purpose of building the intervention.

In the following section, the SMART trial, another type of design that can be used as a stand-alone method or may be useful in the refining phase of MOST, is described. Adaptive interventions and the Sequential Multiple Assignment Randomized Trial (SMART)

In adaptive interventions,⁸ which are also called by other names such as stepped care strategies,^{9,10} treatment algorithms,¹¹ and expert systems,¹² the dose of intervention components may be varied in response to characteristics of the individual or environment. These characteristics are called tailoring variables. The tailoring variable can be something stable like gender or ethnicity, or something that varies over time, such as stage in the Transtheoretical Model,¹² attitude or even progress toward a treatment goal. When the tailoring variable changes over time and there are repeated opportunities to adapt the intervention, this is called a time-varying

adaptive intervention. In adaptive interventions, dosage is assigned to individuals based on a priori decision rules that link values on the tailoring variables to specific intervention dosages. See Collins, Murphy, and Bierman⁸ for a discussion of advantages of adaptive interventions as compared to fixed interventions.

For example, suppose a smoking cessation program includes both positively-framed messages (e.g. “Quitting smoking will help you feel healthier”) and negatively-framed messages (e.g. “Continuing to smoke will increase your risk of serious health problems”). Further suppose that it is expected that those in the precontemplation stage of the Transtheoretical Model are more likely to initiate a quit attempt if presented with a negatively framed message, whereas those in the contemplation stage are more likely to try to quit smoking if presented with a positively framed message. In this example, an individual’s stage in the Transtheoretical Model is the tailoring variable. An adaptive intervention would measure the tailoring variable, i.e. assess whether a smoker is a precontemplator or a contemplator, and deliver a negatively or positively framed message accordingly. A time-varying adaptive intervention would assess this at several different occasions, and once the individual moved from precontemplator to contemplator, would switch to delivering positively framed messages. The strategy “If precontemplator, use negative message framing; if contemplator, use positive message framing” is a decision rule.

The e-health approach lends itself naturally to adaptive interventions. One potential difficulty associated with adaptive interventions is that if decision rules are complex, delivery can be more logistically challenging than that of a comparable fixed intervention. However, a great advantage of e-health is that it can make delivery of even complex time-varying adaptive interventions relatively straightforward. When an e-health approach is used, assessments of tailoring variables can be done electronically, for example, by means of on-line questionnaires and immediate scoring algorithms, and programming algorithms can be used to automate variation of intervention program content or aspects of intervention delivery in response to the tailoring variables. The procedure can be repeated periodically, or as often as each time the individual has contact with the computer program.

BUILDING TIME-VARYING ADAPTIVE INTERVENTIONS USING SMART

SMART is a randomized experimental design that has been developed especially for building time-varying adaptive interventions. Developing an adaptive intervention strategy requires addressing questions such as: What is the best sequencing of intervention components? Which tailoring variables should be used? How frequently, and at what times, should tailoring variables be reassessed and an opportunity for changing dosage presented? Is it better to assign treatments to individuals, or to allow them to choose from a menu of treatment options?

The SMART approach enables the intervention scientist to address questions like these in a holistic yet rigorous manner, taking into account the order in which components are presented rather than considering each component in isolation. A SMART trial provides an empirical basis for selecting appropriate decision rules and tailoring variables. The end goal of the SMART approach is the development of evidence-based adaptive intervention strategies, which are then evaluated in a subsequent RCT.

In a SMART trial, each individual may be randomly assigned to conditions several times. For example, suppose the objective of an intervention is to help people who intend to quit smoking to quit successfully. One question might be, is it better to use positively framed messages or negatively framed messages? Another might be, for individuals who return to smoking, what kind of encouragement to continue to try to refrain from smoking is best, daily email messages or daily email messages augmented with weekly phone calls? For those who are successful at not smoking, is it better to encourage them with daily email messages, or to leave them alone?

And, does the best adaptive intervention strategy vary depending on whether the individual was originally presented with negatively or positively framed messages? A hypothetical SMART trial addressing these questions is outlined in Table 1. At the beginning of the trial, individuals are randomly assigned to receive a DVD with quitting strategies framed positively or framed negatively. At the end of the first month, quitting success is assessed. Those who have not been smoking are randomly assigned to receive either daily email encouragement or no encouragement. Those who have been smoking are given daily email encouragement, and are randomly assigned to receive or not to receive a weekly phone call in addition.

Embedded in this SMART trial are eight different adaptive intervention strategies. These are listed in Table 2. For example, Adaptive Intervention Strategy 3 is, “Begin with a DVD containing smoking cessation strategies framed positively. If at the end of one month the individual is not smoking, no further action is taken. If at the end of one month the individual is smoking, begin sending daily encouraging email messages.” Note that individuals are randomly assigned to adaptive interventions in the SMART approach, but none of the intervention strategies would involve randomization when implemented outside of an experimental setting. In other words, the purpose of the random assignment is to address scientific questions, not to serve as a part of the adaptive intervention.

Suppose the outcome variable in this example is the number of cigarettes smoked during the last week of the study. To address the question “Is it better to use positively or negatively framed messages?” a statistical analysis can be done comparing the mean across adaptive interventions 1+2+3+4 with the mean across adaptive interventions 5+6+7+8. To address the question “For individuals who return to smoking, what kind of encouragement to continue to try to refrain from smoking is best?” the statistical analysis would involve selecting those who respond that they are smoking, and comparing the mean of 1+3+5+7 with the mean of 2+4+6+8. A comparable analysis can be done for those who do not return to smoking. Note that within the SMART trial all of these questions are addressed by means of randomized experiments. Statistical power for each of these analyses is that of a simple two-group comparison. Other scientific questions can be addressed as well. For more about the statistical analysis of SMART trials, see Murphy.¹³ *Integration of SMART and MOST*

Investigators interested in building a time-varying adaptive intervention may find it advantageous to integrate a SMART trial into the MOST procedure. The screening phase of MOST can be used to identify active program components that will be incorporated into the adaptive intervention. The refining phase can be used initially to provide leads for tailoring variables. This can be done by exploring possible interactions between program components and individual and group characteristics. The active program components identified in the screening phase and any tailoring variables identified in the refining phase up to this point can then be used as the basis for a set of time-varying adaptive intervention strategies. The refining phase can continue with a SMART trial to identify the best of these strategies. The confirming phase will then proceed as usual, with a RCT comparing the best adaptive intervention strategy against a suitable comparison group.

CONCLUSION

Because of their reach, e-health interventions promise considerable public health impact. It makes sense to maximize this public health impact by developing the most effective interventions we can. This article has described two related methods for building and evaluating e-health interventions. MOST is an approach for systematically and efficiently optimizing behavioral interventions. The SMART trial is an approach for identifying the best time-varying adaptive intervention strategy. Both approaches are based on randomized experimentation,

which means that a high degree of confidence can be placed in the results. Used individually or together, these methods enable scientists to increase the potency of behavioral interventions.

ACKNOWLEDGEMENTS

This work has been supported by National Institute on Drug Abuse grants P50 DA10075 (Dr. Collins and Dr. Murphy), K05 DA018206 (Dr. Collins), K02 DA15674 (Dr. Murphy), and National Cancer Institute grant P50 CA101451 (Dr. Strecher and Dr. Murphy).

REFERENCES

1. Fox, S. Health information online. 2005. http://207.21.232.103/pdfs/PIP_Healthtopics_May05.pdfhttp://207.21.232.103/pdfs/PIP_Healthtopics_May05.pdf[cited 2007 January 12]; Available from
2. Collins LM, Murphy SA, Nair VN, Strecher V. A strategy for optimizing and evaluating behavioral interventions. *Annals of Behavioral Medicine* 2005;30:65–73. [PubMed: 16097907]
3. Murphy SA. An experimental design for the development of adaptive treatment strategies. *Statistics in Medicine* 2005;24:1455–1481. [PubMed: 15586395]
4. Fisher, RA. The design of experiments. Oliver and Boyd; London: 1935.
5. Box, GEP.; Hunter, WG.; Hunter, JS. Statistics for experimenters: An introduction to design, data analysis, and model building. Wiley; New York: 1978.
6. Fagerlin, A.; Smith, D.; Zikmund-Fisher, BJ.; Derry, H.; Jankovic, A.; Pitsch, R., et al. Society for Medical Decision Making. Cambridge, MA: 2006. Decision aid design influenced patients' knowledge, risk perceptions, and behavior.
7. Wu, CFJ.; Hamada, M. Planning, analysis, and parameter design optimization. Wiley; New York: 2000.
8. Collins LM, Murphy SA, Bierman KL. A conceptual framework for adaptive preventive interventions. *Prevention Science* 2004;5:185–196. [PubMed: 15470938]
9. Breslin FC, Sobell MB, Sobell LC, Cunningham JA, Sdao-Jarvie K, Borson D. Problem drinkers: Evidence of a stepped-care approach. *Journal of Substance Abuse* 1999;10(3):217–232.
10. Sobell MB, Sobell LC. Stepped care as a heuristic approach to the treatment of alcohol problems. *Journal of Consulting and Clinical Psychology* 2000;68(4):573–579. [PubMed: 10965632]
11. Lavori PW, Dawson R, Rush AJ. Flexible treatment strategies in chronic disease: Clinical and research implications. *Biological Psychiatry* 2000;48:605–614. [PubMed: 11018231]
12. Prochaska JO, Velicer WF, Fava JL, Rossi JS, Tsoh JY. Evaluating a population-based recruitment approach and a stage-based expert system intervention for smoking cessation. *Addictive Behaviors* 2001;26:583–602. [PubMed: 11456079]
13. Murphy SA. Optimal dynamic treatment regimes. *Journal of the Royal Statistical Society* 2003;65(2):331–366.Series B

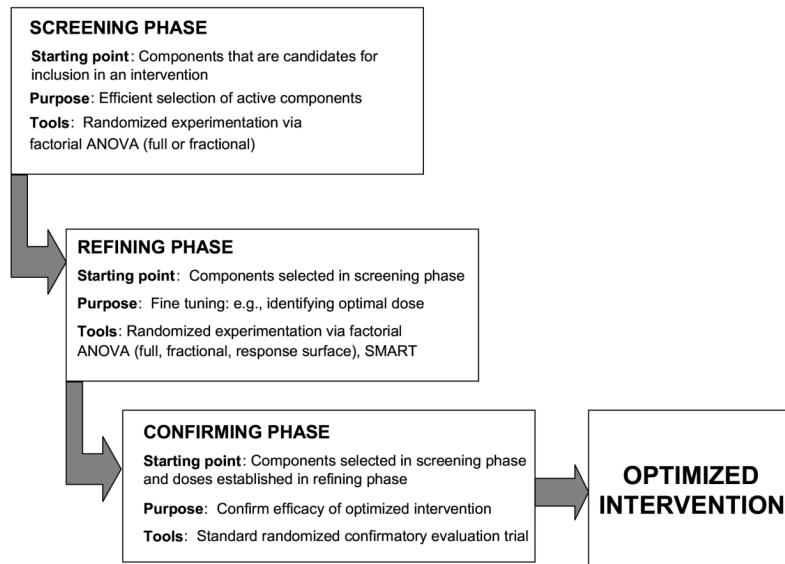


Figure 1. Outline of the Multiphase Optimization Strategy (MOST)
 ANOVA, analysis of variance; SMART, sequential multiple assignment randomized trial

Table 1

Timeline of example SMART trial

Point in trial	Activity	Details
Beginning of trial	Random assignment #1	Each subject randomly assigned to message framing: DVD with positively framed messages OR DVD with negatively framed messages
End of Month 1	Assessment	Asked if they have been smoking
Beginning of Month 2	Random assignment #2	If not smoking at end of Month 1, randomly assigned to: Daily email encouragement OR No encouragement
		If smoking at end of Month 1, randomly assigned to: Daily email encouragement OR Daily email encouragement PLUS weekly phone call
End of Month 2	Assessment	Number of cigarettes smoked in past week assessed

Table 2

Adaptive intervention strategies embedded in SMART trial example in Table 1

Treatment strategy	Framing of messages on DVD	If not smoking at end of Month 1, then:	If smoking at end of Month 1, then:
1	Positive	Daily email encouragement	Daily email encouragement
2			Daily email encouragement PLUS weekly phone call
3		No encouragement	Daily email encouragement
4			Daily email encouragement PLUS weekly phone call
5	Negative	Daily email encouragement	Daily email encouragement
6			Daily email encouragement PLUS weekly phone call
7		No encouragement	Daily email encouragement
8			Daily email encouragement PLUS weekly phone call

SMART, sequential multiple assignment randomized trial