



# HHS Public Access

Author manuscript

*Am J Prev Med.* Author manuscript; available in PMC 2017 May 02.

Published in final edited form as:

*Am J Prev Med.* 2010 January ; 38(1): 98–102. doi:10.1016/j.amepre.2009.09.034.

## Methodologic and Design Issues in Patient-Centered e-Health Research

**Ken Resnicow, PhD, Victor Strecher, PhD, Mick Couper, PhD, Hannah Chua, PhD, Roderick Little, PhD, Vijay Nair, PhD, Thad Polk, PhD, and Audie A. Atienza, PhD**

University of Michigan (Resnicow, Strecher, Couper, Chua, Little, Nair, Polk), Ann Arbor, Michigan; and the National Cancer Institute (Atienza), Bethesda, Maryland

### Introduction

Over the past decade, dozens of studies have demonstrated the impact of e-health interventions across a variety of health domains.<sup>1–6</sup> These include targeted and tailored interventions delivered through various channels such as the web, print, CD/DVD, and more recently PDAs and cell phones. These mostly positive findings have fueled considerable growth in e-health program development. Such programs have tremendous potential for public health impact as they can reach large numbers of participants at a relatively low cost.<sup>7</sup>

Whereas the first generation of patient-centered e-health studies have focused largely on answering the question of *whether* such programs are efficacious; it is proposed that the next generation of e-health research begin to better address the questions of *why*, *how*, and *for whom* they work.<sup>5,8</sup> This brief article discusses some design and analytic approaches for determining what components of e-health programming work (i.e., active ingredients), how they worked (i.e., mechanisms of action), and for whom (i.e., moderator effects).

### Did It Work? What Worked?

Let us begin with an example of a hypothetical e-health Internet-based program designed to help patients quit smoking. The program uses three primary motivational components or factors: outcome expectation messages focusing on reasons for quitting and consequences of continued smoking (Factor A); efficacy expectation messages addressing confidence and barriers to quitting (Factor B); and a narrative testimonial of a patient's specific reasons for, and methods of, quitting (Factor C). This example is a simplified version of an e-health intervention recently developed and evaluated by a number of the authors.<sup>9</sup>

Based on prior empirical studies and theory, factors a, b, and c are all considered important to the outcome of smoking cessation (Y). Most existing e-health interventions have typically examined the aggregate effect of these factors (denoted as abc) in comparison to a control

---

Address correspondence and reprint requests to: Ken Resnicow, PhD, University of Michigan, School of Public Health, Department of Health Behavior and Health Education, 109 Observatory Street, Room 3867 SPH I, Ann Arbor, MI 48109-2029. kresnic@umich.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.

treatment (denoted as 0). Yet, an obvious limitation of such a design lies in the inability to tease out specific contribution of each intervention factor or possible interaction of factors.

The doubly randomized preference trial represents a research design where patients are randomly assigned to a “randomized arm” (where treatments are randomized) and a “choice arm” where patients choose their treatment combination to potentially disentangle intervention factors. However, this approach can yield unmatched control group members with differential exposure to message dose and content. Other options include additive and factorial designs. For example, in the factorial design, the sample of subjects can be randomly divided into eight equal sized groups, and then assign one group to each of the eight treatment combinations, 0, A, B, C, AB, AC, BC, and ABC. Three-way ANOVA can be applied to assess the effects of each factor and their potential interactions. The random allocation to the eight groups ensures that the comparisons among the various ingredients have a valid causal interpretation. Factorial experiments are useful in assessing the main effects of the individual factors A, B, and C, as well as all two- and higher-way interactions.

Full factorial experiments have several drawbacks. A practical difficulty is that the number of treatment combinations (i.e., cells) increases rapidly with the number of factors. For example, if two additional factors d and e could be added to the three-component smoking study, 32 treatment combinations would be required, even with just two levels (e.g., presence/absence) of each factor. A second issue is that there are a large number of statistical comparisons, increasing the number of Type I errors. Additionally, they are less efficient in comparing any two particular treatment combinations, say, ABC against 0. For example, in the original “kitchen sink” approach, ABC and 0 are each assigned  $\frac{1}{2}$  of the subjects while only  $\frac{1}{8}$  of the subjects are assigned in the full factorial experiment with 8 combinations. This reduces the power in comparing specific treatment combinations. In other words, a larger sample size is needed in order to simultaneously assess the existence of an overall effect while also identifying which ingredients and their combinations are most effective.

To address these and other issues, the current research group has begun to use fractional factorial designs to systematically dismantle active ingredients of e-health interventions. These designs were used within a sequential framework which identifies the intervention effects and optimizes their levels before comparing the “optimal” treatment combination against some reference group. Collins et al.<sup>10</sup> call this a multiphase optimization strategy or MOST. This strategy is adapted from a similar framework that has been successfully used in engineering applications for many years (10) and consists of three phases involving separate randomized trials:

**Phase I. Screening**—The goal in this phase is to efficiently “screen” a larger set of potentially important treatment components and identify those with promise. Two-level fractional factorial designs, judiciously chosen to estimate selected higher-order interactions of interest, allow a reduction in the number of experimental groups required. The Pareto principle underlies the screening phase— identifying a relatively small number of active treatment components that account for the greatest variance in the outcome. Since the goal here is to identify potentially important effects, liberal levels were used for significance testing and less consideration is given to multiple comparisons or spurious effects.

**Phase II. Refining**—This phase aims to refine current understanding of the effects of the “candidate” components that are identified in Phase I. This is done using follow-up experiments to untangle important effects that may have been “aliased” in Phase I (i.e., groups of intervention combinations that were not isolated) and to determine optimal “dosage” levels of factors. Phase II results lead to the formulation of an optimal combination of treatment components which can be tested in a fully powered Phase III trial.

**Phase III. Confirming**—The final phase is a confirmation trial to examine the efficacy of the aggregate program built in Phases I and II. While this phase is similar to classic RCTs with two or more arms, the multiphase approach leads to the inclusion of only important components at their optimized levels.

To date, this approach has been successfully implemented in two e-health studies: Project Quit to study the active components of a smoking cessation program<sup>9</sup> and Guide to Decide to study decision aids among women at high risk for breast cancer.<sup>11</sup> Since most behavioral and decision-making interventions in clinical settings rightfully involve multiple treatment components, the MOST approach can serve as an efficient method for “unpacking” the efficacy of each component, thereby maximizing the impact of the overall intervention. A discussion of the use of fractional factorial designs in this context is available elsewhere.<sup>12</sup>

### Limitations of fractional factorial designs

In designing fractional studies, investigators are required to identify, a priori, the key interactions that are likely to be ‘active’. This leads to the aliasing of various other combinations of intervention components, so these interaction effects cannot be isolated. The decisions on which interactions are active and which are inert, while ideally based on prior empirical studies, are often instead based on theoretic assumptions as well as research priorities delineated by the investigators. Potentially important interactions that might not have been anticipated may be excluded from the model, so it is possible to miss small to moderate effects.

Similarly, given that the effect sizes of interactions are rarely known, power estimates are largely hypothetical. This can lead to either underpowered or overpowered studies. Finally, similar to full factorial designs, fractional approaches also involve multiple comparisons and the possibility of Type I error. These limitations can be offset in part by the multiphase confirmatory approach used in the MOST framework whereby even borderline significant effects identified in Phase II are tested again with a new sample and intervention combinations omitted in Phase I can be tested in Phase II.

### How Did It Work?

**Program use data (paradata)**—One of the powerful features of e-health interventions is that they lend themselves to highly specified process analyses. In nonelectronic interventions, process measurements are often limited to constructs such as self-reported use, and user satisfaction (e.g., “Did you read all of the materials?”); however, in most e-health applications, more granular-level details regarding *what* a participant did and *how* they interacted with the materials can be ascertained. These data on intervention use, which

elsewhere have been called “paradata”<sup>13,14</sup>, can be used to help identify on the micro-level which elements of a e-health intervention are used, and to identify participants who were more or less engaged with the materials. This information can then be used to guide the development and refinement of the intervention, subject to the caveat that paradata is observational and subject to the limitations of observational data for making causal inferences.

Several types of paradata can be captured from online interventions. For example in terms of technical specifications, information can be gathered, such as participant browser type, connection speed, and available plug-ins, that may be useful in evaluating feasibility and usability issues. In terms of intervention dose and exposure, information can be obtained regarding number of log-ins, which pages are viewed and for how long, which links are clicked, which applications are launched, and so on. At a finer level of detail, client-side scripts<sup>14</sup> can be run to measure actions within a web page, such as clicking on different objects, scrolling, order of selection of survey response options, among other actions.

In addition to elucidating the consumer uptake of intervention components these data can serve as indicators of engagement which can then be used in “dose–response” analyses.<sup>15</sup> As Danaher et al.<sup>16</sup> note, “a key ingredient in determining the impact of any web-based behavior change program is the extent to which participants are exposed to the program.” Such measures are useful mediators in understanding the effectiveness of online programs. For example, the HMO-based Project Quit program consisted of 5 sections (The program included: an introduction section, a section focusing on outcome expectations, 2 sections focusing on efficacy expectations, and a section with a narrative success story); it was found that each additional section opened by the participant contributed to an 18% higher likelihood of quitting, on average.<sup>17</sup>

**Message processing data**—One method of better understanding how participants process e-health information is to study how they read and attend to the materials. Recent advances in biometrics have expanded the ability to objectively measure, in real time, how materials and messages are processed. For example, tracking eye movements is commonly used in the fields of psychology, engineering, and marketing. People generally have limited awareness and control of how their eyes move under normal viewing conditions. More importantly, eye movements reflect real-time allocation of attention. Eye tracking collects information about patterns of eye movements by providing information about where an individual looks and the duration of their fixations. Fixation positions are important because regions that fall in the eye's fovea are likely to be encoded in greater detail than peripheral regions. Information that is encoded in greater detail is more likely to be remembered later. Some eye tracking devices also collect information about blinks as well as pupil dilation.

Recently, the research group for the current study completed a study examining how testimonials are read as a function of testimonial images that appeared with them.<sup>18</sup> Cigarette smokers were asked to read testimonials of ex-smokers that were paired with images either matched to the participant's age, race, and gender, mismatched to the participant, or neutral with respect to the match. Participants reported that they found the testimonial messages more persuasive when it was paired with a *matched* image. Moreover,

eye-tracking data showed that the smokers viewed the matched image more times, and made more fixations between the matched image and text as if connecting the text and the image. Thus the eye movements appeared to support the persuasion data.

Eye tracking is also useful for testing hypotheses.<sup>19</sup> For instance, suppose a patient is given information about calculating BMI as well as messages about eating fruits. The patient remembers 1 month later reading about the BMI calculation but not the fruit-related messages. In situations like this, it is unclear whether the memory difference was due to a reporting bias or to differences in perception, encoding, or memory consolidation. Eye movement data could potentially distinguish among these hypotheses. For example, if the patient made more eye fixations to BMI calculation than to information about fruits, then the memory difference was probably driven by attention/encoding differences. If the number and duration of fixations were the same in the two conditions, then it is likely that the memory difference was driven by processes that occur after information encoding.

Another promising method for understanding how e-health messages achieve their effects is through functional magnetic resonance or fMRI. fMRI has existed for over 1 decade and has been increasingly utilized in investigating neural and cognitive mechanisms underlying human behavior. This technique can be used to study how e-health messages are processed by patients and how these messages might influence health-related behavior. fMRI investigates neural activity non-invasively and with relatively good spatial and temporal resolution. Participants lie in an MRI scanner and perform tasks while the scanner collects images of blood-oxygen level–dependent (BOLD) signals throughout the brain every few seconds. Because blood-oxygen levels are strongly correlated with neural activity, fMRI can identify brain areas in which neural activity changes as independent variables are changed.

Better understanding patients' neural processing of e-health messages can be used as converging evidence to support existing e-health findings and related theories. Recent research demonstrates that personal relevance appears to partially mediate the effect of tailored messages in smoking abstinence.<sup>17,20</sup> A recent study examined how smokers process high- versus low-depth tailored smoking cessation messages.<sup>18</sup> It was found that smokers activated areas of the brain related to self-relevance processing more when they were processing high-depth tailored messages. Neuroimaging studies can also offer insights into how individuals process information. For example, if other studies find evidence that patients activate a region of the brain when they are engaged in reward processing, and the same area of the brain is activated when a patient is reading a gain-framed motivational health message, it suggests that reward processing may be involved in processing that type of message. Finally, neuroimaging studies are useful in understanding differential cognitive processing based on characteristics of the patient. These characteristics could be genetic, physiologic, behavioral, or psychosocial and could be the basis for new factors used in message tailoring, which is discussed in the next section.

### **For Whom Did It Work and Why?**

Whereas the previous sections have focused on methodologies for determining the effect of specific intervention factors and mechanisms by which they work, it is also important to determine for whom they work and why. One of the most powerful advantages of e-health

interventions is the ability to tailor messages and programs to individual characteristics.<sup>5</sup> Traditionally this task falls under the rubric of moderator analyses (e.g., examining baseline participant characteristics that interact with treatment condition). Typically, in RCTs, moderators include variables such as race, gender, age, and possibly health indicators. However, in e-health interventions, moderators extend beyond these sociodemographic variables. This is in part enabled by the fact that in order to delivery highly tailored interventions, detailed assessments of demographic and psychosocial variables are required. Several tailored and e-health programs<sup>21–23</sup> (including smoking interventions<sup>1,8,17,20,24</sup>) have been shown to interact with a wide range of personality and motivational variables.

One interpretation of these complex interaction patterns is that individual response to e-health interventions may be “sensitive to initial conditions<sup>25,26</sup>.” That is, small initial individual differences (i.e., participant characteristics) may generate large differences in intervention response. If so, this has significant implications for the design, analysis, and replicability of such programs. Although interventions are generally evaluated by comparing overall group differences to a comparison/control group, there may be multiple response patterns that have to be elucidated to fully understand why the intervention worked and for whom. For example, spontaneous quitting may yield more sustained cessation than planned quit attempts<sup>27</sup> and transformative motivation may be more effective than rational motivation.<sup>28</sup>

In addition to executing standard interaction effects, other techniques are available to identify individuals and subgroups of individuals for whom an e-health intervention may be particularly effective.<sup>29</sup> One technique is Signal Detection Methodology (SDM).<sup>29–32</sup> This nonparametric method, which uses Receiver Operator Curves (ROC) and recursive partitioning, identifies homogeneous subgroups (with distinct sets of characteristics) who are higher versus lower on an outcome parameter (e.g., success with a smoking cessation e-health intervention). In contrast to standard interaction effects which tend to look at one or two levels, SDM can convey higher-order interactions, providing a potentially richer characterization of those for whom an intervention was successful and not. Incorporating a broader range of traditional moderator variables, as well as paradata, in SDM analyses, can enable e-health researchers to identify for whom interventions are more and less effective and ultimately to create more effectively tailored interventions.

The more frequent interactions patients can have with e-health programs, often entailing repeated assessment of psychosocial and behavioral processes (occasionally in real time), lend themselves to more highly powered and granular analyses of *why* behavior change may occur. Thus, the ability to examine potential mediators or mechanisms of behavior change can be significantly enhanced with fine-grained, time-stamped data obtained in e-health programs.<sup>33</sup> Taken together, through both designing e-health interventions to assess theoretic mechanisms and examine paradata, researchers can develop a better understanding of the “black box” of why and how these programs work.

## Conclusion

This article proposes some novel design and analytic approaches to determining how, why, and for whom aggregate and subcomponents of patient-centered e-health program work. Each of the approaches to addressing the questions described in this article has been used by the current research group as well as by a number of others.<sup>8</sup> A better understanding of what is inside the black box of e-health interventions will lead to more empirically informed and ultimately more effective programs. Furthermore, innovative analytic techniques should be employed to determine for whom e-health interventions are most effective. The refinement of current and future e-health interventions promises to increase the effectiveness of these interventions for specific populations and future research employing the analytic strategies discussed above is warranted to examine this possibility.

## References

1. Webb MS, Simmons VN, Brandon TH. Tailored interventions for motivating smoking cessation: using placebo tailoring to examine the influence of expectancies and personalization. *Health Psychology*. 2005; 24(2):179–88. [PubMed: 15755232]
2. Noar SM, Clark A, Cole C, Lustria ML. Review of interactive safer sex websites: practice and potential. *Health Communication*. 2006; 20(3):233–41. [PubMed: 17137415]
3. Suggs LS. A 10-year retrospective of research in new technologies for health communication. *Journal of Health Communication*. 2006; 11(1):61–74. [PubMed: 16546919]
4. Noar SM, Benac CN, Harris MS. Does tailoring matter? Meta-analytic review of tailored print health behavior change interventions. *Psychological Bulletin*. 2007; 133(4):673–93. [PubMed: 17592961]
5. Hawkins R, Kreuter M, Resnicow K, Fishbein M, Dijkstra A. Understanding Tailoring in Communicating About Health. *Health Education Research*. 2008; 23(3):454–66. [PubMed: 18349033]
6. Yap TL, Davis LS. Physical activity: the science of health promotion through tailored messages. *Rehabilitation Nursing*. 2008; 33(2):55–62. [PubMed: 18330383]
7. Strecher V. Internet methods for delivering behavioral and health-related interventions (e-health). *Annual Review of Clinical Psychology*. 2007; 3:53–76.
8. Dijkstra A. Working mechanisms of computer-tailored health education: evidence from smoking cessation. *Health Education Research*. 2005; 20(5):527–39. [PubMed: 15701665]
9. Strecher VJ, McClure JB, Alexander GL, Chakraborty B, Nair VN, Konkel JM, et al. Web-based smoking-cessation programs: results of a randomized trial. *American Journal of Preventive Medicine*. 2008; 34(5):373–81. [PubMed: 18407003]
10. Box, GEP., Hunter, WG., Hunter, JS. *Statistics for Experimenters*. New York: Wiley; 1978.
11. Zikmund-Fisher BJ, Ubel PA, Smith DM, Derry HA, McClure JB, Stark A, et al. Communicating side effect risks in a tamoxifen prophylaxis decision aid: the debiasing influence of pictographs. *Patient Education and Counseling*. 2008; 73(2):209–14. [PubMed: 18602242]
12. Nair V, Strecher V, Fagerlin A, Ubel P, Resnicow K, Murphy S, et al. Screening experiments and the use of fractional factorial designs in behavioral intervention research. *American Journal of Public Health*. 2008; 98(8):1354–9. [PubMed: 18556602]
13. Couper, MP., Lyberg, LE. 55th Session of the International Statistical Institute; 2005. Sydney, Australia: 2005. *The Use of Paradata in Survey Research*.
14. Heerwegh D. Explaining response latencies and changing answers using client-side paradata from a web survey. *Soc Sci Comput Rev*. 2003; 21(3):360–373.
15. Glasgow RE, Nelson CC, Kearney KA, Reid R, Ritzwoller DP, Strecher VJ, et al. Reach, engagement, and retention in an Internet-based weight loss program in a multi-site RCT. *Journal of Medical Internet Research*. 2007; 9(2):e11. [PubMed: 17513282]

16. Danaher BG, Boles SM, Akers L, Gordon JS, Severson HH. Defining participant exposure measures in web-based health behavior change programs. *Journal of Medical Internet Research*. 2006; 8(3):e15. [PubMed: 16954125]
17. Strecher VJ, McClure J, Alexander G, Chakraborty B, Nair V, Konkel J, et al. The role of engagement in a tailored web-based smoking cessation program: RCT. *Journal of Medical Internet Research*. 2008; 10(5):e36. [PubMed: 18984557]
18. Chua HF, Liberzon I, Welsh RC, Strecher VJ. Neural Correlates of Message Tailoring and Self-Relatedness in Smoking Cessation Programming. *Biological Psychiatry*. 2009; 65(2):165–168. [PubMed: 18926523]
19. Chua, HF., Boland, JE., Nisbett, RE. Cultural variation in eye movements during scene perception. *Proceedings of the National Academy of Sciences of the U. S.*; 2005; p. 12629-33.
20. Strecher VJ, Shiffman S, West R. Moderators and mediators of a web-based computer-tailored smoking cessation program among nicotine patch users. *Nicotine & Tobacco Research*. 2006; 8(1):S95–101. [PubMed: 17491176]
21. Resnicow K, Davis RE, Zhang G, Konkel J, Strecher VJ, Shaikh AR, et al. Tailoring a fruit and vegetable intervention on novel motivational constructs: results of a randomized study. *Annals of Behavioral Medicine*. 2008; 35(2):159–69. [PubMed: 18401673]
22. Gans KM, Risica PM, Strolla LO, Fournier L, Kirtania U, Upegui D, et al. Effectiveness of different methods for delivering tailored nutrition education to low income, ethnically diverse adults. *International Journal of Behavioral Nutrition and Physical Activity*. 2009; 6:24. [PubMed: 19416525]
23. Resnicow K, Davis R, Zhang N, Tolsma D, Alexander G, Wiese C, et al. Tailoring a fruit and vegetable intervention on ethnic identity: Results of a randomized study. *Health Psychol*. 2009; 28(4):394–403. [PubMed: 19594262]
24. Webb MS, Hendricks PS, Brandon TH. Expectancy priming of smoking cessation messages enhances the placebo effect of tailored interventions. *Health Psychology*. 2007; 26(5):598–609. [PubMed: 17845111]
25. Resnicow K, Vaughan R. A chaotic view of behavior change: a quantum leap for health promotion. *International Journal of Behavioral Nutrition and Physical Activity*. 2006; 3(1):25. [PubMed: 16968551]
26. Resnicow K, Page SE. Embracing chaos and complexity: a quantum change for public health. *American Journal of Public Health*. 2008; 98(8):1382–9. [PubMed: 18556599]
27. West R, Sohal T. “Catastrophic” pathways to smoking cessation: findings from national survey. *BMJ*. 2006; 332(7539):458–60. [PubMed: 16443610]
28. Matzger H, Kaskutas LA, Weisner C. Reasons for drinking less and their relationship to sustained remission from problem drinking. *Addiction*. 2005; 100(11):1637–46. see comment. [PubMed: 16277625]
29. King AC, Ahn DF, Atienza AA, Kraemer HC. Exploring refinements in targeted behavioral medicine intervention to advance public health. *Annals of Behavioral Medicine*. 2008; 35(3):251–60. [PubMed: 18568380]
30. Kiernan M, Kraemer HC, Winkleby MA, King AC, Taylor CB. Do logistic regression and signal detection identify different subgroups at risk? Implications for the design of tailored interventions. *Psychological Methods*. 2001; 6(1):35–48. [PubMed: 11285811]
31. King AC, Kiernan M, Oman RF, Kraemer HC, Hull M, Ahn D. Can we identify who will adhere to long-term physical activity? Signal detection methodology as a potential aid to clinical decision making. *Health Psychology*. 1997; 16(4):380–9. [PubMed: 9237091]
32. Kraemer, HC. *Evaluating medical tests: Objective and quantitative guidelines*. Newbery Park: Sage; 1992.
33. Weinstein SM, Mermelstein R, Shiffman S, Flay B. Mood variability and cigarette smoking escalation among adolescents. *Psychology of Addictive Behaviors*. 2008; 22(4):504–13. [PubMed: 19071975]