Examining the boundaries of tailoring: the utility of tailoring versus targeting mammography interventions for two distinct populations

G. L. Ryan, C. S. Skinner¹, D. Farrell² and V. L. Champion³

Abstract

Health messages can be generic, targeted to population subsets or tailored for individual recipients. There has been little examination of which populations need tailored interventions or whether tailored and targeted interventions differ in important ways. We used data from a mammography intervention study in two distinct populations to simulate a comparison of individually tailored versus targeted interventions. Tailored intervention content was based on individual recipients’ interview responses. Targeted intervention content was based on composite group responses. For more than 60% in each population group, about two-thirds of tailored message content was a good match with content of the targeted intervention generated by composite group responses; roughly one-third of the content was ‘not a good’ fit for their intervention needs. Tailored interventions for more than 80% of subjects in each population differed in at least some way from those generated for all other population group members. This simulation is a first step in quantifying the contribution of individual tailoring over group targeting. Future research should examine whether a targeted intervention that is mostly a ‘good’ match results in behavioral outcomes similar to those of individually tailored interventions and whether particular differences in tailored versus targeted interventions yield significantly more favorable intervention outcomes.

Introduction

Health education messages can be communicated in a number of ways, from generic waiting room pamphlets providing general information, to one-on-one counseling sessions with messages specifically tailored for each counselee. Mass-produced materials may be generic, targeted to certain population subsets, personalized on identifiers such as name, or individually tailored to a person’s psychosocial or behavioral characteristics (Kreuter et al., 1998). These communication strategies differ by level of assessment of the audience (i.e. what is discovered about audience members before developing messages) and degree of individualization (i.e. whether all audience members receive the same message version or a ‘customized’ or ‘tailored’ version) (Kreuter et al., 1998, 2000a). Some materials are generic in that they require no audience assessment and have no customized content. Targeted materials are specially developed for a population subgroup and require some group-level assessment to focus, or ‘target’, messages for the group. Although the most common targeting may be based on demographic characteristics, messages can also be targeted based on behavioral or psychosocial characteristics of group members. Finally, tailored health messages are based on assessments of individuals and are individually produced for each intervention recipient.
As computer-tailoring technology has developed, mass-produced tailored health communication has emerged as a promising strategy for influencing attitudes and behaviors (Skinner et al., 1994, 1998; Campbell et al., 1994; Strecher et al., 1994). Randomized controlled trials have shown tailored mammography interventions can successfully impact mammography beliefs and behavior (Champion, 1994; Skinner et al., 1994). This success may be explained by the Elaboration Likelihood Model which theorizes that people process information more actively and thoughtfully if they perceive it to be personally relevant (Petty and Cacioppo, 1981). Indeed, studies have shown tailored health messages are better remembered, read and/or perceived as relevant (Skinner, 1994; Campbell et al., 1994; Brinberg and Axelson, 1990; Brug et al., 1996, 1998). However, tailoring research is a relatively new field, leaving many questions to be answered before we understand how to optimally tailor health promotion messages, and how, when and where tailored health communication will be most useful.

Tailoring studies have differed widely by theoretical framework, sample characteristics, study design, message style, behavioral target and outcomes measured (Rakowski, 1999; Skinner et al., 1999). As a result, studies lack a consensus regarding when tailoring is necessary, what amount of detail is required in the individualization and on which variables messages should be tailored. Because tailoring technology theoretically allows an infinite number of sociodemographic, psychosocial, behavioral and clinical variables to be addressed, it will be important to examine the ‘boundaries and limits of tailoring’ (Abrams et al., 1999). As Kreuter et al. comment, ‘being able to generate millions of different message combinations is one thing, needing to do so is another’ (Kreuter et al., 1999).

In examining the boundaries of tailoring, investigators should evaluate audiences and health behavior outcomes for which tailored health interventions are most effective. For example, if a given population is fairly homogeneous with respect to factors influencing mammography use, such as race, income and perceived barriers to mammography screening (Lane and Fine, 1983; Howard, 1987; Kruse and Philips, 1987; Rutledge et al., 1988; Tippy et al., 1989; Wolosin, 1989; Lerman et al., 1990; Champion, 1991; Calnan et al., 1992; Skinner et al., 1994, 1998; Dolan et al., 1995; Burns et al., 1996), individually tailored mammography messages may deliver relatively similar information to most audience members. As Kreuter et al. note, ‘there is little need to tailor a message if it would end up being more or less the same for all members of the intended audience’ (Kreuter et al., 1999). This may be especially true of one-time behaviors (i.e. undergoing a mammogram) as opposed to lifestyle changes (i.e. diet or exercise) with fewer variables affecting the outcome (Kreuter et al., 1998). Our challenge is to develop parsimonious theoretical models outlining what is worth tailoring, for what types of people and in what sociocultural contexts (Abrams et al., 1999).

To address this challenge, we used data from an ongoing study of computer-tailored mammography promotion interventions to simulate what difference might exist between tailored and targeted mammography interventions for two distinct populations. We first assessed variation within each population’s tailored health messages by calculating a ‘rate of uniqueness’—the degree to which each individually tailored message combination generated for each subject by the tailoring program differed from other combinations generated for subjects in the same population. Second, we explored whether, for a majority of each of the two study populations, a well-targeted mammography intervention varied significantly from individually tailored messages (with regard to mammography-related beliefs, stage of adoption and self-efficacy). Finally, to investigate the role demographic characteristics play in predicting variance from a targeted intervention, we explored whether an intervention targeted to the combined populations varied significantly from the individually tailored interventions.
Methods

Study population
Study data were from a comparison of tailored print and telephone communications in two university-affiliated sites—an HMO in Indianapolis and a general medicine clinic in St Louis. Women patients at either site who were at least 51 years old, had no mammography within 15 months, no history of breast cancer and were able to complete telephone interviews were eligible to participate. Procedures were approved by the respective universities’ Human Subjects Committees.

Data collection
Data are from the 20-min baseline telephone interviews and the computer-tailored letter profiles generated for each woman, based on her responses. Interviews assessed breast cancer knowledge, perceived risk, mammography benefits and barriers, self-efficacy for obtaining a mammogram, stage of mammography adoption, cancer fatalism, and demographic characteristics.

Computer-tailored interventions
Tailored newsletters were generated as follows. Algorithms prepared by the investigators determined which sets of baseline interview responses triggered selection of which messages. A program written in FileMaker Pro (version 3.0 1990, 1992–1996 Claris Corp., Santa Clara, CA) read each woman’s interview data, chose relevant message texts from libraries of potential messages and placed the selected message texts at nine pre-determined locations, or ‘tailoring points’, on each tailored cover letter and newsletter.

Two models of health behavior framed the tailored interventions. The Health Belief Model (HBM) suggests that perceived susceptibility to a health threat, severity of that threat, and benefits, barriers, and self-efficacy associated with protecting oneself influence the likelihood of taking that protective action (Becker, 1974; Janz and Becker, 1984; Rosenstock et al., 1988). A number of studies have shown correlations between HBM variables and mammography use (Curry and Emmons, 1994), and several have demonstrated the effectiveness of mammography interventions tailored on HBM constructs (Skinner et al., 1994; Champion and Huster, 1995) The second model—the Transtheoretical Model (TM)—suggests that health behavior change includes a series of stages differentiated by whether an individual has engaged in a health behavior and, if not, what her future intention is (Prochaska et al., 1992). Stages have been defined and validated for mammography adoption (Rakowski et al., 1992, 1993) and used in other tailored mammography interventions (Champion, 1994; Skinner et al., 1994; Champion and Huster, 1995).

In our study, each tailored intervention included a cover letter and a newsletter of 1–3 pages, depending on interview responses. Cover letters, printed on clinic or HMO stationary, included recipient’s name, address and personal physician’s signature. Text included general information (thanks for participation, introduction to newsletter, reference to gender and age as the two main breast cancer risk factors, and a general closing statement), a reference to recipient’s age and the following two ‘tailoring points’:

1. Reference to presence or absence of breast cancer family history.
2. Reference to subject’s mammography stage of adoption.

Newsletters were printed with eye-catching color and graphics. Page 1 text included general information regarding mammography’s benefits and ‘tailoring points’ (3)–(7):

3. Introduction to recipient’s perceived breast cancer risk.
4. Reference to one or two risk factors specifically mentioned by recipient or, in absence of specific risk factors, reference to her perceived risk.
5. Introduction to recipient’s level of perceived benefits of mammography.
6. Reference to benefits not identified by recipient in the interview.
(7) Reference to one, two or three barriers mentioned by recipients in the interview or a general message regarding mammography barriers (included if she mentioned less than two specific barriers).

Newsletter pages 2 and 3 consisted of ‘tailoring points’ (8) and (9):

(8) The ‘procedure page’ with colored icons and text instructing women how to schedule an appointment, what to wear and what to expect during the appointment was included if the recipient was in the contemplation stage (i.e. thinking about having her first mammogram), or indicated difficulty making an appointment, finding the time to go for the mammogram or knowing what to expect.

(9) The ‘self-efficacy page’ with photos of women modeling the behavior of obtaining a mammogram, along with encouraging text was included for recipients with low self-efficacy scores or expectations that screening center staff would be unfriendly.

More than 2 billion possible combinations of tailored messages could be generated for an intervention with these nine tailoring points. This product was calculated by multiplying the number of possible messages at each of the nine ‘tailoring points.’

**Rate of uniqueness**

To assess amount of variation existing among tailored interventions within each population, we calculated a ‘rate of uniqueness’. Equal to the percentage of interventions containing completely unique message combinations (e.g. different from the combination generated for any other recipient), rate of uniqueness indicates how extensively the tailoring program produced unique message combinations, as opposed to generating the same message combinations for multiple recipients.

The ‘rate of uniqueness’ calculation was possible because each available message at each of the nine ‘tailoring points’ had been given a designated number in the computer program. Thus, each tailored intervention was coded by the computer as a combination of nine numbers. Using a statistics program, we ordered the 1163 tailored interventions numerically and scanned for unique combinations.

**Generation of targeted interventions**

The first step in exploring whether a group-targeted mammography intervention would have varied significantly from individually tailored messages for the two distinct populations was to create a well-targeted intervention for each population. Thus, we generated frequency distributions for each interview question and entered the most common responses (the mode) into the computer-tailoring program. The resulting two printed interventions (each consisting of a cover letter and a 1–3 page newsletter) represent targeted interventions for the two populations because each was developed based on a composite population profile rather than characteristics of any specific member of the population (see Table I). We repeated this process with most common responses from the combined populations to develop a third targeted intervention based on a composite profile of the entire study population from both sites (see Table I). Because targeted interventions were not sent to women in the sample, this paper reports on a simulation meant to quantify differences in output using targeting versus tailoring methods rather than an actual intervention trial.

**Generation of ‘match scores’ for each tailoring point**

We next assessed, for each of the two populations and the combined population, whether the targeted interventions varied significantly from individually tailored interventions. The method of comparison was as follows. At each of the nine tailoring points, we compared each possible message in the message library with the particular message generated for the targeted message (by the process described above). Every message in each of the nine libraries received a ‘match score’ of 0, 0.5 or 1 point, based on how well it matched the information generated for that particular point in the targeted intervention.
**Table I. Decision rules for the generation of ‘match scores’ at each ‘tailoring point’**

<table>
<thead>
<tr>
<th>Tailoring point</th>
<th>Message in the hypothetical well-targeted intervention</th>
<th>Messages receiving a ‘match score’ of 1 point</th>
<th>Messages receiving a ‘match score’ of 0.5 points</th>
<th>Messages receiving a ‘match score’ of 0 points</th>
</tr>
</thead>
<tbody>
<tr>
<td>(1) Family history of breast cancer</td>
<td>no</td>
<td>no</td>
<td>no</td>
<td>yes</td>
</tr>
<tr>
<td>(2) TM stage of mammography compliance</td>
<td>relapse</td>
<td>relapse</td>
<td>relapse</td>
<td>any non-relapse stage</td>
</tr>
<tr>
<td>(3) Level of perceived risk of breast cancer</td>
<td>average or lower</td>
<td>average or lower</td>
<td>average or lower</td>
<td>high</td>
</tr>
<tr>
<td>(4) Specific risk factors mentioned by subject</td>
<td>none; low perceived risk</td>
<td>none; low perceived risk</td>
<td>none; low perceived risk</td>
<td>any factors</td>
</tr>
<tr>
<td>(5) Level of perceived benefit of mammography screening</td>
<td>high</td>
<td>high</td>
<td>high</td>
<td>lower</td>
</tr>
<tr>
<td>(6) Specific benefits known to subject</td>
<td>find cancer early found early find cancer early found early</td>
<td>find cancer early found early find cancer early</td>
<td>find cancer early found early find cancer early</td>
<td>any other</td>
</tr>
<tr>
<td>(7) Specific barriers to compliance</td>
<td>discomfort</td>
<td>no barrier</td>
<td>no barrier</td>
<td>any other barrier</td>
</tr>
<tr>
<td>(8) Procedure page included with instructions and what to expect</td>
<td>no</td>
<td>no</td>
<td>no</td>
<td>yes</td>
</tr>
<tr>
<td>(9) Self-efficacy page included with encouraging photos and text</td>
<td>no</td>
<td>no</td>
<td>no</td>
<td>yes</td>
</tr>
</tbody>
</table>
We did this at each of the nine tailoring points using criteria as follows (also see Table I). At most tailoring points, if the message in the targeted and tailored interventions matched exactly, we assigned a ‘match score’ of 1. If different messages were included in the targeted and tailored interventions, a 0 was assigned. For example (see Table I), at ‘tailoring point’ (7), the message in the targeted intervention for the Indianapolis population addressed discomfort associated with mammography (the most commonly identified barrier among this population). Thus, the discomfort message was given a ‘match score’ of 1; the other 18 possible messages in the library for this tailoring point received a 0 ‘match score.’ This dichotomous value system was adopted based on the goal of tailoring—that an individual’s messages in her intervention refer to her specific needs and characteristics.

At ‘tailoring points’ (2) and (4), however, we adopted a more conservative strategy based on the text message. We did this when the text selected for a subject by the computer-tailoring program was not identical to the message in the targeted intervention, yet was similar to it. We felt it unfair to rate these particular messages as complete mismatches (with a score of 0 points) since doing so would have underestimated the degree of ‘fit’ between the message content and recipient’s needs and thus artificially inflated the difference between the targeted and tailored messages. For example, at the tailoring point for mammography stage ['tailoring point’ (2)] the targeted intervention for the St Louis population addressed relapse from adherence with mammography recommendations (the most common stage in this population). We assigned this relapse message a value of 1 point and most other messages in the library (those addressing contemplation, precontemplation, action or maintenance) a value of 0 points because they truly did not fit the needs of the recipient. However, one message in the message library at this tailoring point (addressing those at risk for relapse) was assigned a value of 0.5 points because its content was similar to the message in the targeted intervention. It was not a complete mismatch. Decision rules for each message library were discussed and agreed upon by principal investigators of the larger study and by another researcher active in the field of tailored health interventions, and are displayed in Table I.

**Generation of ‘match score sums’ for each tailored intervention**

After each possible message in each of the nine libraries was compared to the message generated for the well-targeted letter and given a ‘match score’ of 0, 0.5 or 1, a ‘match score sum’ was calculated for each subject’s individually tailored intervention. ‘Match score sums’ indicate overall fit of the individual tailored letter with the targeted letter and, thus, directly addresses the study’s examination of the boundaries of tailoring mammography interventions.

The ‘match score sum’ represents how well each individually tailored intervention ‘matches’ the population’s targeted intervention. A tailored intervention received a maximum sum of 9 points if its messages were identical to the population’s well-targeted letter at all nine ‘tailoring points’ and a ‘match score sum’ of 0 points if it did not match or resemble the targeted intervention at any tailoring point.

Placing these ‘match scores’ into meaningful categories, we deemed tailored interventions with ‘match score sums’ of at least 6 (of 9) points a ‘good match’ with the targeted letter; 3 to 5.5 points were a ‘fair match; and fewer than 3 points were a ‘poor match.’ This quantitative grouping system was constructed by investigators to give some workable value to the ‘match score sums.’ Absent a gold standard for what is a good match, it was logical to construct the groups by dividing the maximum ‘match score sum’ into thirds. Although this categorization does not address overall qualitative ‘match’, the grouping provides a good base for discussion and further investigation of outcomes based on these categories.

**Statistical comparison of the populations**

Using the above-mentioned ‘match score sums’, we calculated percentages of tailored interventions
with quantitatively ‘good’, ‘fair’ and ‘poor’ matches with the targeted intervention in each of the two populations. Bivariate analyses compared these percentages between the Indianapolis and St Louis populations. Finally, we used the composite targeted intervention—generated by mode responses from combined Indianapolis and St Louis samples—to calculate percentage of tailored interventions that were quantitative ‘good’, ‘fair’ and ‘poor’ matches for the more heterogeneous combined-sample group. Results were statistically compared to the ‘good’, ‘fair’ and ‘poor’ percentages in each of the two separate populations.

Results

Study population
The 501 Indianapolis subjects were members of the Indiana University Medical Group, a large managed-care primary practice administered and staffed by attending physicians from Indiana University’s Division of General Internal Medicine. Most women were Caucasian (80%), high school graduates (81%), not currently employed (55%) and had children (89%). Mean age was 65 years (SD = 11 years), mean years of education was 13 (SD = 3 years) and mean total yearly income was $30–50 000. The 662 St Louis subjects were patients in the General Medicine Clinic of Barnes-Jewish Hospital, an outpatient clinic located in the hospital and staffed by residents and attending physicians from Washington University’s Division of General Medical Sciences. Most subjects at this site were African-American (83%), not currently employed (91%) and had children (86%). Mean age was 68 years (SD = 10 years), mean education was 11 years (SD = 3 years) and mean total yearly income was $15–30 000.

In addition to site of healthcare delivery, each population was homogeneous in age, race, parity and number of first-degree relatives with breast cancer. The St Louis population was also homogeneous in its income, employment and marital status (see Table II). Significant differences between the two populations were found in age, race, marital status, education, employment status and income ($P < 0.001; see Table II).

Rate of uniqueness
Eighty-two percent of tailored interventions delivered to women in the Indianapolis population were completely unique in their message combinations. In St Louis, 85% of the tailored interventions were unique from all others developed for members of this population.

Match score sums
A majority of tailored interventions in each population received a ‘match score sum’ of at least 6 out of 9 possible points (see Table III). Of tailored interventions generated for the Indianapolis population, 69% were a quantitatively ‘good’ match with the Indianapolis targeted intervention. In other words, 69% of the individually tailored interventions received at least 6 ‘match score’ points based on how well each message in the intervention matched or resembled the corresponding message in the Indianapolis targeted intervention. Twenty-nine percent of the tailored interventions in this population were a quantitatively ‘fair’ match for the targeted intervention, receiving at least 3, but fewer than 6, ‘match score’ points. Only 1% were a quantitatively ‘poor’ match, receiving fewer than 3 ‘match score’ points.

Of tailored interventions generated for the St Louis population, 61% were a quantitatively ‘good’ match with the targeted St Louis intervention (receiving at least 6 ‘match score’ points), 37% were a ‘fair’ match and 2% were a quantitatively ‘poor’ match.

When compared to the well-targeted intervention for the entire study group, a majority of tailored interventions for the St Louis and Indianapolis populations also received at least 6 ‘match score’ points. Specifically, 65% of tailored interventions generated for the composite population were a quantitatively ‘good’ match (for the intervention targeted to the entire study population from both sites), 33% were a quantitatively ‘fair’ match and 2% of the tailored interventions were a quantitatively ‘poor’ match.
Table II. Demographic indicators of the Indianapolis and St Louis populations

<table>
<thead>
<tr>
<th>Demographics</th>
<th>Indianapolis</th>
<th>St Louis</th>
<th>Significance</th>
</tr>
</thead>
<tbody>
<tr>
<td>Age</td>
<td>mean: 64.6</td>
<td>mean: 67.8</td>
<td>$P &lt; 0.01$</td>
</tr>
<tr>
<td></td>
<td>SD: 10.9</td>
<td>SD: 9.9</td>
<td></td>
</tr>
<tr>
<td></td>
<td>range: 51–97</td>
<td>range: 51–92</td>
<td></td>
</tr>
<tr>
<td>Race (%)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Caucasian</td>
<td>80</td>
<td>15</td>
<td></td>
</tr>
<tr>
<td>non-Caucasian</td>
<td>20</td>
<td>85</td>
<td></td>
</tr>
<tr>
<td>Marital status (%)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>with partner</td>
<td>42</td>
<td>18</td>
<td></td>
</tr>
<tr>
<td>with no partner</td>
<td>58</td>
<td>82</td>
<td></td>
</tr>
<tr>
<td>Education (%)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>&lt;high school</td>
<td>19</td>
<td>53</td>
<td></td>
</tr>
<tr>
<td>high school graduate</td>
<td>37</td>
<td>26</td>
<td></td>
</tr>
<tr>
<td>&gt;high school</td>
<td>44</td>
<td>20</td>
<td></td>
</tr>
<tr>
<td>Employment status (%)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>not employed</td>
<td>55</td>
<td>91</td>
<td></td>
</tr>
<tr>
<td>employed part-time</td>
<td>12</td>
<td>5</td>
<td></td>
</tr>
<tr>
<td>employed full-time</td>
<td>33</td>
<td>4</td>
<td></td>
</tr>
<tr>
<td>Yearly income [$ (%)]</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>&lt;15 000</td>
<td>27</td>
<td>78</td>
<td></td>
</tr>
<tr>
<td>15–30 000</td>
<td>33</td>
<td>11</td>
<td></td>
</tr>
<tr>
<td>30–50 000</td>
<td>21</td>
<td>3</td>
<td></td>
</tr>
<tr>
<td>50–75 000</td>
<td>10</td>
<td>0.3</td>
<td></td>
</tr>
<tr>
<td>75–100 000</td>
<td>4</td>
<td>0</td>
<td></td>
</tr>
<tr>
<td>100–150 000</td>
<td>2</td>
<td>0.2</td>
<td></td>
</tr>
<tr>
<td>No. of first degree relatives with breast cancer (%)</td>
<td>87</td>
<td>88</td>
<td></td>
</tr>
<tr>
<td></td>
<td>12</td>
<td>10</td>
<td></td>
</tr>
<tr>
<td></td>
<td>1</td>
<td>1</td>
<td></td>
</tr>
<tr>
<td></td>
<td>0</td>
<td>0.2</td>
<td></td>
</tr>
<tr>
<td>Have children (%)</td>
<td>89</td>
<td>86</td>
<td>$\chi^2 = 3; P = 0.05$</td>
</tr>
</tbody>
</table>

Statistical analysis of ‘match score sum’ distributions showed no significant difference between percentages of tailored interventions in the Indianapolis population that were ‘good’, ‘fair’ and ‘poor’ matches for the Indianapolis targeted intervention, and percentages of tailored interventions in the St Louis population that were ‘good’, ‘fair’ and ‘poor’ matches for the St Louis targeted intervention. This was also true when distributions for each of the two populations were compared individually to the distribution for the composite study population.

**Discussion**

This investigation was prompted by observations among project staff who noticed they repeatedly heard the same responses to interview questions and/or saw similar message combinations produced by the tailoring program. Given the large number of possible tailored message combinations—more than 2 billion—we wondered whether this study had passed the point of diminishing returns in creating such a complex list of tailoring variables and such an extensive list of message options. This question became especially relevant given the apparent demographic homogeneity of the subjects in each population. Would it be just as effective, and more efficient, to target a mammography intervention to each population subgroup rather than individually tailoring the content of each message to recipients’ knowledge of breast cancer, perceived susceptibility, benefits, barriers and self-
Examining the boundaries of tailoring

**Table III. Match score sums for the separate and combined populations**

<table>
<thead>
<tr>
<th>Match score sums</th>
<th>Indianapolis (n = 501) (%)</th>
<th>St Louis (n = 662) (%)</th>
<th>Composite group (n = 1163) (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>6–9 = a ‘good’ match:</strong> this proportion of tailored interventions received a total of at least 6 ‘match score’ points based on how well each message exactly matched or resembled their population’s hypothetical well-targeted intervention at each of the nine tailoring points</td>
<td>69</td>
<td>61</td>
<td>65</td>
</tr>
<tr>
<td><strong>3–5.5 = a ‘fair’ match:</strong> this proportion of tailored interventions received a total of at least 3, but less than 6, points based on comparison with their population’s well-targeted intervention</td>
<td>29</td>
<td>37</td>
<td>33</td>
</tr>
<tr>
<td><strong>0–2.5 = a ‘poor’ match:</strong> this proportion of tailored interventions received a total of less than 3 points based on this process of comparison</td>
<td>1</td>
<td>2</td>
<td>2</td>
</tr>
</tbody>
</table>

efficacy, stage of mammography adoption, cancer and fatalism, and demographic characteristics? Without the help of established methodology for quantitatively comparing tailored and targeted interventions, the methods described in this paper were an attempt to begin to answer this question by way of process analysis and simulation.

The ‘rate of uniqueness’ measure gave an initial rough indication of how extensively the computer-tailoring program was being used to produce unique messages rather than repeatedly generating the same message combinations. Results of this measurement showed 82–85% of tailored interventions for the two populations were completely unique from all the other interventions tailored for other subjects in the same population.

Whereas the absolute number of message combinations generated by the computer program in these two populations is only a fraction of what the program is capable of generating, the calculated rates of uniqueness still suggest the vast majority of subjects received a tailored message combination not shared by any other recipient in their city. Looking at this result another way, a well-targeted intervention sent to all members of each population would be off-base at one or more ‘tailoring points’ for at least 80% of subjects in each population. This suggests that, within the context of this study’s specific questionnaire and computerized tailoring program, different messages were generated for at least some of the nine ‘tailoring points.’ However, it may be that these differences were at only a few, and not all, of the 9 ‘tailoring points.’

To more closely compare group-targeted versus individually tailored interventions, we created a targeted intervention and assigned values to messages in the tailoring library based on how each compared, qualitatively, to the corresponding message in the targeted intervention. We used this system to evaluate each tailored intervention, thus simulating a comparison of tailored versus targeted interventions. Despite the high rate of uniqueness in each population, results showed the majority of tailored interventions was a quantitatively ‘good’ match for the targeted intervention produced for the population. This was true when the individual interventions for both Indianapolis and St Louis were compared to the targeted intervention for each corresponding population. Thus, although a small minority of the individually tailored interventions were a perfect fit with the targeted intervention and few were exactly the same as other individually tailored interventions, a majority of the mass-produced tailored interventions were at least two-thirds identical in content as the targeted intervention for each population.

By comparing separately for each population and then together as a combined group, it is possible to see what contribution demographic variability may have on predicting how well
targeted interventions may approximate tailored mammography interventions. Interestingly, the proportion of individually tailored interventions for Indianapolis and St Louis that were a ‘good’ match with the targeted intervention produced for the composite group was remarkably similar to the percentages in the comparisons of the separate populations. In other words, this simulation showed that the intervention specifically targeted to one particular subgroup (e.g. the mostly low-income African-American group in St Louis or the mostly Caucasian HMO members in Indianapolis) did not match that subgroup much better than the intervention targeted to the combined group of non-adherent women over 50 years of age.

It may be that, insofar as these variables could be evaluated by this study’s baseline interview and addressed by this study’s tailoring program, all women over age 50 and non-adherent for mammography are relatively similar with respect to their breast cancer knowledge, perceived risk, benefits, barriers, and self-efficacy associated with mammography, stage of mammography adoption and cancer fatalism.

Our findings are, of course, affected by our decision to treat each behavioral characteristic and each ‘tailoring point’ as similarly important. That is, each of the nine ‘tailoring points’ was given equal value in evaluating a quantitative ‘match’ between the tailored and targeted interventions. However, matching a message to a particular characteristic may be more important for some subjects or on some variables than others. For instance, specifically addressing one woman’s stage of mammography adoption may be more important than specifically reinforcing her perceived benefits. For another intervention recipient, the opposite may be true. It may be that targeted interventions that were quantitatively ‘good’ matches would be qualitatively ‘fair’ or ‘poor’ matches for the individually tailored interventions if the different ‘tailoring points’ were given different weights in evaluating a match.

Clearly, the decision to categorize ‘match score sums’ of 6–9 as a quantitatively ‘good’ match, 3–5.5 as a ‘fair’ match and 0–2.5 as a ‘poor’ match influenced the results of the study. If we had decided to use ‘match score sums’ of 7–9 as a ‘good’ match, for example, a minority (specifically 32–41%) of tailored interventions would have been categorized as quantitatively ‘good’ matches with the targeted interventions. Any change in the system of categorization may have also changed whether there was a significant difference in ‘matching’ between the separate populations and the entire, composite population. Because this paper describes a mere first step toward answering the many questions concerning tailored interventions, there remain many possible techniques for future analysis. For example, instead of assessing the degree to which each tailored output message produced by the program matches the corresponding message segment, a targeted intervention date could be factor analyzed to produced models of several distinctive sets of messages and investigators could look at total matches for the entire output rather than using the message-by-message assessment adopted for each tailoring point.

Despite the lack of accepted methodology and various possibilities for comparison, we believe the available data support our heuristic categorization. A recent report asked a similar question using non-tailored versus tailored interventions (Kreuter et al., 2000b). Subjects in the non-tailored group whose information exactly matched the tailored information they would have received (had they been in the tailored group) on at least 70% of the 17 variables in the intervention had outcomes not significantly different from those who did receive tailored interventions, whereas non-tailored recipients whose designated information matched on less than 70% of the variables had significantly worse outcomes than those in the tailored group. These findings underline the importance of investigating necessary and sufficient amounts of ‘matching’ in a health behavior intervention. The results also support our decision to categorize a ‘poor’ match as less than a 33% match, a ‘fair’ match as at least 33% but less than 67% match and a ‘good’ match as at least 67% match. In conclusion, our analysis revealed that despite demographic homogeneity within the two populations, more
than 80% of each population received a unique tailored intervention. Furthermore, demographics did not predict the outcome of the comparison between individually tailored interventions and group-targeted interventions—the targeted intervention was a ‘good’ match for more than 60% of the individually tailored interventions regardless of demographics. Clearly, this simulation investigation is one of many needed to further elucidate how best to use the tool of tailoring in health communications. Further study should investigate whether a well-targeted intervention that is a ‘good’ match results in a similar behavioral outcome to that of a tailored intervention or, alternately, whether the minority of content that did not match the targeted output was crucial for behavior change. Perhaps our goal should be to ask a few important questions of intervention recipients in order to develop an optimally individualized, efficiently mass-produced health behavior intervention individually tailored on key variables but, on non-crucial variables, targeted to a wider audience.

**Acknowledgements**

This investigation was funded by grant no. PHS R01 NR04081 from the National Institute of Nursing Research. The authors wish to thank Tamara Crowe, Tiffany Tibbs, Julia Hobberger, Linda Gidday, Helen Todora and Judy Musick at the Washington University School of Medicine, Matthew W. Kreuter at the St Louis University School of Public Health, and Usha Menon at the Indiana University School of Nursing for their invaluable assistance in conducting this study and preparing this report.

**References**


G. L. Ryan et al.


Received on March 8, 1999; accepted on April 24, 2000