

Increasing Preventive Health Services via Tailored Health Communications

Kathleen T. Durant, PhD; Jack Newsom, ScD; Elizabeth Rubin, MPA; Jan Berger, MD, MJ; and Glenn Pomerantz, MD

In recent years, the focus on healthcare quality has increased, but insurers face simultaneous pressures to control costs. Efforts to improve quality need to be directed in a measurable and cost-effective manner. This paper describes a method for developing health communication outreaches that addresses healthcare quality in a resource-limited scenario.

The US Preventive Services Task Force (USPSTF) reviews the evidence of the effectiveness of clinical preventive measures and develops recommendations for clinical preventive services.¹ These clinical recommendations are incorporated into the definition of the Healthcare Effectiveness Data and Information Set (HEDIS) measures. Implementing these recommendations has been shown to prevent and detect disease early, when it is most treatable and curable.¹

The USPSTF's current recommendation for colorectal cancer screening (CRCS) is for men and women of average risk to begin screening at age 50 years and continue through age 75 years. The screenings can be implemented as an annual high-sensitivity fecal occult blood test (FOBT); a colonoscopy every 10 years; or a flexible sigmoidoscopy every 5 years along with a high-sensitivity FOBT every 3 years.¹ Early detection of high-risk lesions, polyps, or CRC itself through screenings is associated with decreased incidence of, and mortality from, CRC.

Despite the proven benefits of CRCS, only 58.6% of the targeted US population completes a CRCS.² This is well below the national target of 70.5%,² and may explain why it is the third-most common cancer in the United States, as well as the second-leading cause of cancer mortality.³ Total direct cost associated with treating CRC is \$17 billion annually.^{3,4}

Historically, physicians communicated about preventive care, such as CRCS, to the patient. More recently, as health plans increasingly take on risk associated with quality and outcomes, plans have augmented this physician-led activity with their own communications, since health communications have been shown to be beneficial insofar as they

ABSTRACT

Objectives

To develop a methodology that stratifies members by likelihood of completing a colorectal cancer screening (CRCS). Such information can guide the communication development and the allocation of resources for tailored communication outreaches.

Study Design

Prospective study of an insured commercial population that includes randomized assignments to a control group or a treatment group for communication interventions.

Methods

A total of 46,697 members overdue (nonadherent based on available administrative data) for a CRCS were assigned to 1 of 5 graded segments as part of an interactive voice response (IVR) call. Another 400 members per outreach segment were randomly assigned to a control group and received no communication. Primary outcome: receipt of CRCS (assessed Healthcare Effectiveness Data and Information Set [HEDIS] summary data, 3 months after IVR) for each graded segment within the control and the intervention groups. Secondary outcomes: association between the CRCS rates and the identified segments; communication cost per completed screening.

Results

Primary outcome: 1971 members completed the screening. Screening rates for the 5 graded segments were 2.4%, 3.6%, 5.0%, 7.2%, and 8.0%. Secondary outcome: χ^2 proportions test a significant association between the segment groups and the CRCS completion rates. The association between segment and screening rates follow the expected predicted range in both the intervention and holdout groups. Communication cost per screening: \$14.84.

Conclusions

Segmenting members by variables that are associated with CRCS completion identified graded segments that completed the CRCSs at different rates. Tailored interventions can be developed to promote a health service for each segment. The segmentation approach proves to be beneficial.

Am J Manag Care. 2014;20(10):828-835

increase the audience's knowledge and support of the topic, refuting misconceptions, and ultimately leading to the desired behavior.⁵

Typically, however, insurers' communication outreaches are homogeneous. In fact, heterogeneity will increase the success of these outreaches. It is important to: design a tailored outreach that will resonate with the recipient (member); choose a channel the member is likely to use; deliver the message when the member is most receptive to it; and allot ample time for completion before the end of the year. This paper describes a methodology that allows development and measurement of such a communication outreach. The importance of this methodology is not only in the CRCS completion rates, but also in the proper allocation of funds for plans with scarce resources.

METHODS

The study was reviewed by the Horizon Blue Cross Blue Shield of New Jersey (BCBSNJ) privacy official and business owner, both whom deemed the study to fall under "healthcare operations," not "research," meaning it did not require approval by the Internal Review Board.

The **Figure** displays the process flow for the study as well as the different populations (segmented population vs not segmented population) and the treatment groups (intervention group vs control group).

Decision Tree Creation

We gathered health, healthcare utilization, and demographic variables from 2010 and 2011 claims data for 97,821 members of an insurance plan from Horizon BCBSNJ eligible for CRCS. A majority of the decision tree variables were identified as relevant population characteristics of people completing CRCS in prior research.⁶⁻³⁷ As suggested by prior research, we included a member's past health behavior to predict fulfillment of a healthcare service in the future.³⁸⁻⁴¹

We used a modification of IBM Statistical Package for the Social Sciences' (SPSS) χ^2 automatic interaction detection growing algorithm (CHAID SPSS Predictive Analytic Software (PASW) statistics version 18)⁴² to build decision trees that identify the most relevant variables in CRCS completion. Decision trees separate a collection of heterogeneous records into smaller groups of homogeneous records using a directed knowledge discovery. During each step of the tree-growing process, the CHAID algorithm

Take-Away Points

- Readily available quantitative data derived from claims, demographic attributes, and past Healthcare Effectiveness Data and Information Set activity can be used to segment a population based on their likelihood to fulfill a particular health service.
- The segmentation scheme can guide the development of a predictive outreach that supports the personalization of communications in order to improve the quality of healthcare services.
- This communication model can be used to improve the allocation of communications-oriented resources.

chooses the independent variable that has the strongest interaction with the dependent variable for the population specific to a tree leaf node.⁴² We considered other machine-learning techniques, but chose to use decision trees since their nonlinear nature allows tree induction to exploit structures that identify subpopulations discerned by the tree variables.⁴²

Given the distinct variables associated with gender, we chose to develop separate models for men (45,678) and women (52,143). We used 2 different validation techniques: 10-fold cross-validation and a validation set consisting of 33% of the members; the validation methods verify the chosen variables, as well as the final results from the model.

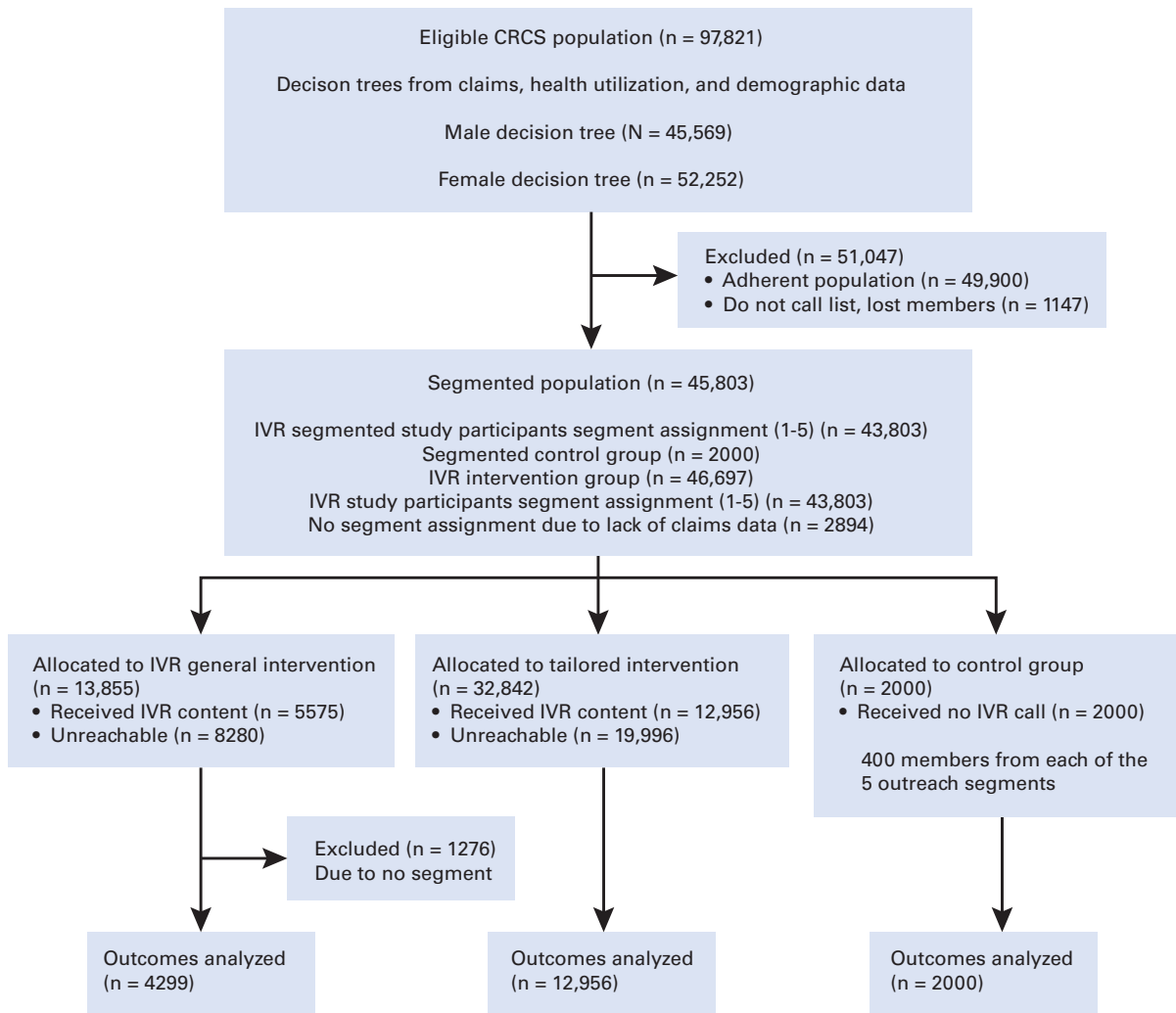
Trees to Segments

To segment members, we split the distribution of the predicted class probability of the outcome variable (CRCS completion = 1). The predicted probability is determined for each leaf node during the model creation step and is defined as the proportion of cases in each leaf node that complete the screening. We use the class probability values to define the segments, since previous studies have shown that decision trees are known for estimating accurate probability-based rankings.⁴³ The segment cut points within the probability distribution maximize the difference in probability values among the segments, as well as create segments that are large enough to be actionable. The goal is to make the members of a segment as homogeneous as possible while making the segments heterogeneous.

Outreach Definition

The outreach was defined as 1 interactive voice response (IVR) call with varying messaging. Depending on the number of nonadherent members and the segments' health descriptions, an outreach segment may contain 1 or more model segments. **Table 1** shows the percentage of male and female populations assigned to each outreach segment.

■ **Figure.** Member Process Flow



CRCS indicates colorectal cancer screening; IVR, interactive voice response.

Power Analysis

A power analysis was performed before the launch of the intervention to determine the minimal size needed for each segment, given an estimated effect size of 2% increase for each graded segment and an α level = .05. Given that the sizes of the segmented groups were determined by the insurer’s nonadherent population and not via a recruitment method, it was determined at the launch of the communication that the comparison of the completion rate of segment 4 and segment 5 was underpowered given the segment sizes and the estimated effect size.

Outcomes Analysis

We measured the predictive nature of the methodology by measuring the CRCS completion rates and calculating the 95% CI for each of the 5 segments for the intent-

to-treat population and the 2 groups (holdout group and treatment group). The study’s segmented population is defined as the segmented population that were targeted to receive either the general intervention, the tailored intervention or are part of the control group. We also compare the results of the control group to the reached intervention group. The control group is defined as the segmented population that was targeted to not receive the IVR call. The reached intervention group are members that actually received the intervention (the IVR content). These members were reached by telephone and may have received either the targeted content or the general content. To measure the stability of the segments, we compared each segment’s completion rates for the 2 different treatment groups. Given the different proportions of the graded segments within the holdout and intervention groups, we did not

compare the completion rates of the 2 groups; we hypothesized that the holdout group would have a higher completion rate since it has a higher proportion of members from the higher-ranked segments. A demographic comparison of the 2 groups' segments can be found in the eAppendix tables.

RESULTS

1971 members completed the screening with 1873 screening completions from members assigned an outreach segment (1773 from the intervention group and 100 from the control group) and 98 screening completions from members not assigned a segment due to a lack of claims data for the calendar years 2010 and 2011. Segments predicted to have a higher CRCS completion rate did follow the predicted range in both groups. The range of segment completion rates for the study's segmented population, from lowest to highest probability, were 2.42%, 3.64%, 4.83%, 7.06%, and 8.08%.

Decision Tree Variables

For females, other cancer screenings such as breast cancer screening, as well as a personal or family history of colon cancer, are the most predictive variables and make up the topmost levels of the decision tree. These variables have been previously identified as associated with CRCS.^{7,10,13,18,26,28,29,37} Another variable in the tree, but not previously identified, is total claims spending. Other relevant variables included having an obstetrics/gynecological visit, completing a cervical cancer screening,^{12,13,26} completing a physical exam,^{17,18,26} the number of doctor visits,^{14,17-19,22,30,35} and a radiology exam. A member's age and being coded as having unexplained abdominal pain or musculoskeletal chronic pain are also included within the female decision tree.

For men, who, as specified by the old HEDIS specification, do not have other cancer-related HEDIS measures, the most predictive variable is claims spending; this variable has not been identified as associated with CRCS in previous studies. Other variables such as a personal or family history of CRC and the number of doctor visits for the prior 2 calendar years (2010, 2011) are variables that make up the topmost levels of the decision tree. These variables have been identified as associated with

Table 1. Description of the Groups

	Control Group n = 2000	Intervention Group n = 43,803
Gender male	894 (44.70%)	17,663 (48.54%)
Female	1106 (55.30%)	18,725 (51.46%)
Age mean	$\mu = 58.1$	$\mu = 58.2$
Standard deviation	$\sigma = 5.4$	$\sigma = 5.3$
CRCS model segment		
1	399 (19.95%)	14,618 (33.37%)
2	314 (15.70%)	11,164 (25.49%)
3	752 (37.60%)	11,268 (25.72%)
4	472 (23.60%)	5838 (13.33%)
5	63 (3.15%)	915 (2.09%)
CRCS indicates colorectal cancer screening.		

CRCS in previous studies.^{7,10,17,19,20,29,30,37} Having a physical exam^{17,18,26} and a variable indicating a household's net worth^{7,10,17,30,37} are other variables within the male decision tree. Household net worth is typically correlated to income and socioeconomic class, variables identified as associated with CRCS completions in previous studies.^{10,17,30,37} Other variables within the decision tree are the number of emergency department (ED) visits and unexplained abdominal pain.

Predictability of the CHAID Trees

Table 2 lists the area under the receiver operating characteristics (ROC) curve (AUC) for the decision trees. We list the percentage of the majority class as an indication of the predictive power of a model. AUC in the range presented are consistent with those associated with other models used to predict healthcare behavior, notably models designed to identify members at risk of hospital readmission.⁴⁴

Segment Definition

Table 2 lists the identified segments, along with each segment's percentage of the total population and the percentage overdue for a CRCS before the outreach. The segments are graded such that a higher number indicates a segment more likely to complete the CRCS. Table 2 shows the models are predictive and the segments have varying CRCS completion rates before the outreach. For example, roughly 1 out of 20 men in Segment 5 is overdue for a CRCS, while 16 out of 20 men in Segment 1 are overdue for a CRCS.

IVR Outreach

A total of 46,697 members overdue for a CRCS received

Table 2. Predictability of the Tree Models and the Derived Segments From the Trees

Gender-Specific Colorectal Cancer Screening Completion Models	Area Under the Curve (AUC)		AUC 95% CI		Majority Percentage	
	Male	Female	Male	Female	Male	Female
	0.653	0.702	[0.648-0.658]	[0.697-0.707]	50.7%	52.5%

Colorectal Segments	Segment Percentage of the Population per Gender		Percent of the Segment Overdue for a CRCS Before IVR Outreach	
	Male	Female	Male	Female
Segment Rank 5	7.4%	10.9%	5.70%	15.16%
Segment Rank 4	17.0%	28.0%	35.51%	32.00%
Segment Rank 3	33.0%	22.9%	53.65%	47.58%
Segment Rank 2	22.1%	17.3%	64.15%	69.99%
Segment Rank 1	20.5%	21.0%	79.14%	84.80%
Total	100%	100%	54.55%	51.35%

CRCS indicates colorectal cancer screening; IVR, interactive voice response

Table 3. Actual and (Expected) Counts and CRCS Completion Rates for the Study’s Segmented Population 90 Days Post Treatment

CRCS	Study Population				
	Segment 1	Segment 2	Segment 3	Segment 4	Segment 5
Completed (expected)	367 (621.1)	446 (500.7)	535 (453.0)	446 (258.2)	79 (40.0)
Not completed (expected)	14,821 (14,566.9)	11,798 (11,743.3)	10,543 (10,625.0)	5869 (6056.8)	899 (938.0)
Completion rate	2.42% [2.18%, 2.68%]	3.64% [3.02%, 3.68%]	4.83% [4.60%, 5.38%]	7.06% [6.43%, 7.69%]	8.08% [6.37%, 9.79%]
Total	15,188	12,244	11,078	6315	978

CRCS indicates colorectal cancer screening.

an IVR communication call between January 23, 2013, and February 5, 2013, encouraging the screening. The outreach made up to 2 attempts to reach the member. Calls resulting in a busy or no-answer result were retried up to 3 times. Each outreach segment, defined by the decision tress, had 400 members that were part of the control group, for a total of 2000 members in the holdout group. Forty percent of the intervention group received the IVR call.

Statistical Analysis

We performed a Pearson χ^2 analysis using the CRCS completion as the dependent variable and the ordinal segment identifier as an independent variable for the study's segmented population, the holdout group, and the intervention group. The tests show that there is an association between the segment ranking and the CRCS completion ($P < .0001$, the study's segmented population; $P < .0001$, intervention group; $P = .0118$, control group). **Tables 3** and **4** contain the actual, expected completed, and not completed CRCS counts for the 5 segments;

completion rates for the holdout and the intervention groups are similar. However, there is overlap in some of the adjacent segments' 95% CI for the completion rates in the intervention and holdout groups. For the the study's segmented population segments' results, it is only the top 2 more adherent segments that contribute relatively fewer members to the outreach where we see an overlap in the CIs. This finding is in alignment with the low-power level for the χ^2 statistic.

Communication Cost-Effectiveness

The communication cost was \$29,317, and it yielded 1971 members completing at least 1 type of CRCS, leading to a net communication cost of \$14.84 per screening. We can further refine the analysis by taking into account the different time lengths of CRC prevention provided by the screenings. A colonoscopy provides 10 years of CRC prevention while an FOBT provides 1 year. Given that out of the 1971 colorectal screenings, 1863 were colonoscopies and 108 were FOBTs, the number of CRC pre-

■ **Table 4.** CRCS Completion Rates and Actual and (Expected) Counts for the Reached Intervention Group Versus the Control Group

Segments	Completion Rates ± CI Rate Range, 95% CI		[done, not done] Actual (done, not done) Expected	
	Treatment	Control	Treatment	Control
Segment 1	3.71% ± .54% [3.17%-4.25%]	2.0% ± 1.37% [0.63%-3.37%]	[176, 4566] (264.8, 4477.2)	[8, 392] (20, 379.1)
Segment 2	4.93% ± .74% [4.19%-5.39%]	4.79% ± 2.36% [2.43%-7.16%]	[238, 4590] (269.6, 4558.4)	[15, 298] (15.7, 298.3)
Segment 3	6.10% ± .71% [5.39%-6.81%]	5.32% ± 1.6% [3.72%-6.92%]	[270, 4157] (247.2, 4179.8)	[40, 712] (37.6, 714.4)
Segment 4	8.26% ± 1.01% [7.15%-9.15%]	6.57% ± 2.24% [4.33%-8.80%]	[229, 2577] (156.7, 2649.3)	[31, 441] (23.6, 448.4)
Segment 5	11.02% ± 2.85% [8.17%-13.87%]	9.52% ± 7.24% [2.28%-16.77%]	[51, 412] (25.9, 437.1)	[6, 57] (3.2, 59.9)

CRCS indicates colorectal cancer screening.

vention years is the number of completed colonoscopies multiplied by 10 plus the number of completed FOBTs multiplied by 1. This calculates to 18,738 CRC prevention years, with a resultant communication cost of \$1.56 per CRC prevention year.

DISCUSSION

Preventive care services, such as CRCS, have been shown to improve the quality of healthcare by increasing the detection of disease at earlier stages, when treatment is more likely to lead to a cure. The CDC has shown that although CRCS is widely recognized as important based on evidence-based medicine, only 58% of patients in the target group are completing a screening.¹ Over the last several years, insurers have created outreach activities to attempt to have more of their members complete CRCS. One challenge that insurers face is how to best attempt these outreach activities in an environment of limited resources. It therefore becomes imperative to identify outreach methods that improve the use of these services in a cost-effective manner.

This paper presents a methodology that identifies graded segments for CRCS completion, which then can become the basis for an insurer's communication plan. Table 3 shows that the least compliant segment completes CRCS at 0.6x the average completion rate for the total population of the study, while the highest compliant segment completes CRCS at 1.98 times the average completion rate of the study. This variation in predicted completion rates can guide the scheduling and number of outreaches as well as the percentage of allocated resources.

For women, variables representing preventive care are the most influential variables in the model. These vari-

ables include breast cancer screening adherence, cervical cancer screening adherence, visits to an obstetrics/gynecology office, doctor office visits, and completion of a physical exam.

For men, monetary variables are the most influential variables, with the most predictive variable being higher claims spending. Men having lower than average claims spending and not having a doctor's office visit in 2 years will not likely be CRCS-adherent; interestingly, the predicted probability associated with the adherence increases as the approximated monetary worth of the household increases. A variable associated with lower CRCS rates is ED visits. In reviewing the other variables associated with this subpopulation, they seem to be reactive in healthcare and to use the ED as their primary source of care. Identifying a subpopulation's approach to healthcare allows us to craft relevant content for that segment.

For both genders, having a personal or family history with CRC is a strong signal for CRCS completion. However, this variable affects a relatively small proportion of the population. Variables associated with CRCS completion in both genders and not identified in other studies are higher claims spending and unexplained abdominal pain. These findings may be specific to this particular population or plan design. The association between unexplained abdominal pain and CRCS may be due to a health provider using 1 of the accepted screenings as a diagnostic tool to investigate the abdominal pain. Interestingly, not all types of pain led to higher CRCS rates in a female subpopulation. Other types of pain such as back, joint, or musculoskeletal pain led to slightly lower CRCS rates in a population adherent in breast cancer screening yet ineligible for cervical cancer screening due to aging out of the denominator (women over 65 years of age). This associa-

tion between lower CRCS rates and musculoskeletal pain should be reassessed in another demographically similar female population.

Within this study we monitor the CRCS completion of 5 segments, where members are assigned to a holdout group (no health message sent) or an intervention group (1 IVR call). A χ^2 proportions test comparing the CRCS rates for the 5 graded segments shows that the CRCS rates do differ by segment (the study's population group, $P < .0001$; the reached intervention group, $P < .0001$; holdout group, $P = .0018$) but each segment has similar rates given different treatments. We believe the segments' similar rates in the 2 different groups is due to factors such as the small size of each segment's holdout group, the stability of the segments, and the homogeneity of the treatment. Segments' CRCS rate variation can be exploited when developing a health communication outreach. In particular, the health outreach can be varied by calendar timing; number of communications; type of outreach (short message service [SMS], IVR, letter, agent call); and content. We believe varying these outreach levers will provide an increase in CRCS completion rates, leading to reduced inequities in healthcare participation ultimately leading to a reduction in health inequities.

Limitations

The quality of the analysis is dependent on the quality of the claims coding or any other source of the models' variables. The health variables were derived from claims data and may only reflect the billing of health services and not accurately reflect the health or the healthcare utilization of the member. Within this study, there was little variation in plan design; this may account for the identified association between CRCS and total claims spending.

The methodology was applied to 1 East Coast commercial population and may not generalize to other populations and other health services beyond CRCS. This intervention defined the number of outreaches, the calendar timing, and the mode of the outreach to be the same for all 5 segments. This limited the utility of the segmentation scheme but allowed us to investigate the segments' stability given a small deviation in the treatment delivered to the groups.

We used the segments to develop communication interventions for a nonadherent population; the segments may be applicable for developing other personalized areas of healthcare and used to formulate a successful communications strategy. The methodology's utility should be further explored in future studies.

CONCLUSION

We present a methodology that segments a population by likelihood to complete a preventive care screening. The methodology can be applied to develop a communication intervention where the outreach can be varied by calendar timing, number of communications, type of outreach, and the content of the outreach. We have created 1 outreach using this methodology where only 1 out of 4 levers varied. The results from the initial study show the methodology is predictive and is promising for developing communication outreaches. However, it is early in the process to evaluate the methodology's total benefits, since we have not taken advantage of all communication levers. We plan to take a stepwise approach to outreach development, since this will allow us to iteratively measure and identify changes that lead to increases in the outcome variable (CRCS). Overall, the segmentation approach proves to be beneficial, allowing for the proper allocation of limited resources for the best possible outcomes. The targeted approach for segmented populations demonstrates that communications can drive higher adherence rates, especially with limited resources.

Acknowledgments

We would like to acknowledge and thank several Horizon Blue Cross Blue Shield New Jersey employees for their contributions to this project. Lori Leotta, director of Clinical Initiatives, has been a supporter of this project. Brian Prestipino, senior informatics analyst, and John DiNatale, senior informatics manager, reviewed and prepared the members' claims data that were part of the intervention. We also would like to thank Carmela Uzzi, senior writer, and Melissa Rioseco, communications business partner, for their review of an earlier draft of the manuscript.

Author Affiliations: Silverlink Communications, Burlington, MA (KTD, JN, JB); Horizon Blue Cross Blue Shield of New Jersey, Newark, NJ (ER, GP).

Source of Funding: Methodology developed through funding from Silverlink Communications. Horizon Blue Cross Blue Shield of New Jersey funded Silverlink Communications to develop the colorectal cancer screening outreach to be sent to their eligible commercial population.

Author Disclosures: Drs Durant and Newsom are employees of Silverlink Communications, have attended meetings and conferences for the company, and own stock options. Dr Berger is an employee of Silverlink Communications. Dr Pomerantz is an employee of Horizon Blue Cross Blue Shield of New Jersey. Ms Rubin has no financial interests to disclose.

Authorship Information: Concept and design (KTD, JN, JB, GP); acquisition of data (KTD, JN, ER); analysis and interpretation of data (KTD, GP); drafting of the manuscript (KTD); critical revision of the manuscript for important intellectual content (KTD, JN, ER, JB); statistical analysis (KTD); provision of study materials or patients (ER); obtaining funding (JN); administrative, technical, or logistic support (JN, ER, JB, GP); supervision (JN, JB).

Address correspondence to: Kathleen T. Durant, PhD, Silverlink Communications, 1 Burlington Business Center, 67 South Bedford St, Suite 300, Burlington, MA 01803. E-mail: kdurant@silverlink.com.

REFERENCES

1. US Preventive Services Task Force. Screening for colorectal cancer: US Preventive Services Task Force recommendation statement. <http://www.uspreventiveservicestaskforce.org/>. *Ann of Intern Med*. 2008;149(9):627-637.
2. HealthyPeople.gov. Healthy People 2020: Topics & Objectives: Cancer: Objective C-16. <http://healthypeople.gov/2020/topics/objectives/2020/objectiveslist.aspx?TopicId=5>.
3. Jemal A, Tiwari RC, Murray T, et al; American Cancer Society. Cancer statistics, 2004. *CA Cancer J Clin*. 2004;54(1):8-29.
4. American Cancer Society. What are the key statistics about colorectal cancer? <http://www.cancer.org/cancer/colonandrectumcancer/detailedguide/colorectal-cancer-key-statistics>.
5. National Cancer Institute. *Making Health Communication Programs Work: A Planner's Guide, revised edition*. Bethesda, MD: National Cancer Institute; 2001.
6. Mosen DM, Feldstein AC, Perrin NA, et al. More comprehensive discussion of CRC screening associated with higher screening. *Am J Manag Care*. 2013;19(4):265-271.
7. Gimeno Garcia AZ. Factors influencing colorectal cancer screening participation. *Gastroenterol Res Pract*. 2012. 2012:doi10.1155/2012/483417.
8. Semrad TJ, Tancredi DJ, Baldwin LM, Green P, Fenton JJ. Geographic variation of racial/ethnic disparities in colorectal cancer testing among Medicare enrollees. *Cancer*. 2011;117(8):1755-1763.
9. Blom J, Yin L, Lidén A, et al. Toward understanding nonparticipation in sigmoidoscopy screening for colorectal cancer. *Int J Cancer*. 2008;122(7):1618-1623.
10. Neilson AR, Whynes DK. Determinants of persistent compliance with screening for colorectal cancer. *Soc Sci Med*. 1995;41(3):365-374.
11. Wei N, Nakama H, Li T. Screen compliance rates in 14-years annual screening program for colorectal cancer with immunochemical fecal occult blood test—identification of higher priority subjects in health education. *Eur J Med Res*. 1998;3(7):341-344.
12. Dear K, Scott L, Chambers S, Corbett MC, Taupin D. Perception of colorectal cancer risk does not enhance participation in screening. *Therap Adv Gastroenterol*. 2008;1(3):157-167.
13. Shapiro JA, Seeff LC, Nadel MR. Colorectal cancer-screening tests and associated health behaviors. *Am J Prev Med*. 2001;21(2):132-137.
14. López-Charneco M, Pérez CM, Soto-Salgado M, et al. Correlates of colorectal cancer screening among Hispanics: Results from the 2008 Puerto Rico behavioral risk factor surveillance system survey. *P R Health Sci J*. 2013;32(2):68-75.
15. Castañeda SF, Xiong Y, Gallo LC, et al. Colorectal cancer educational intervention targeting Latino patients attending a community health center. *J Prim Care Community Health*. 2012;3(3):164-169.
16. McVeigh TP, Lowery AJ, Waldron RM, Mahmood A, Barry K. Assessing awareness of colorectal cancer symptoms and screening in a peripheral colorectal surgical unit: a survey based study. *BMC Surg*. 2013;13:20.
17. Daly JM, Xu Y, Levy BT. Patients whose physicians recommend colonoscopy and those who follow through. *J Prim Care Community Health*. 2013;4(2):83-94.
18. McQueen A, Vernon SW, Meissner HI, Klabunde CN, Rakowski W. Are there gender differences in colorectal cancer test use prevalence and correlates? *Cancer Epidemiol Biomarkers Prev*. 2006;15(4):782-791.
19. Seeff LC, Nadel MR, Klabunde CN, et al. Patterns and predictors of colorectal cancer test use in the adult US population. *Cancer*. 2004;100(10):2093-2103.
20. Slattery ML, Kinney AY, Levin TR. Factors associated with colorectal cancer screening in a population-based study: the impact of gender, health care source, and time. *Prev Med*. 2004;38(3):276-283.
21. Subramanian S, Klosterman M, Amonkar MM, Hunt TL. Adherence with colorectal cancer screening guidelines: a review. *Prev Med*. 2004;38(5):536-550.
22. Ioannou GN, Chapko MK, Dominitz JA. Predictors of colorectal cancer screening participation in the United States. *Am J Gastroenterol*. 2003;98(9):2082-2091.
23. Price JH. Perceptions of colorectal cancer in a socioeconomically disadvantaged population. *J Community Health*. 1993;18(6):347-362.
24. Beeker C, Kraft JM, Southwell BG, Jorgensen CM. Colorectal cancer screening in older men and women: qualitative research findings and implications for intervention. *J Community Health*. 2000;25(3):263-278.
25. Wardle J, Miles A, Atkin W. Gender differences in utilization of colorectal cancer screening. *J Med Screen*. 2005;12(1):20-27.
26. Brawarsky P, Brooks DR, Mucci LA. Correlates of colorectal cancer testing in Massachusetts men and women. *Prev Med*. 2003;36(6):659-668.
27. Weinberg DS, Miller S, Rodoletz M, et al. Colorectal cancer knowledge is not associated with screening compliance or intention. *J Cancer Educ*. 2009;24(3):225-232.
28. Weinberg DS, Turner BJ, Wang H, Myers RE, Miller S. A survey of women regarding factors affecting colorectal cancer screening compliance. *Prev Med*. 2004;38(6):669-675.
29. Stockwell DH, Woo P, Jacobson BC, et al. Determinants of colorectal cancer screening in women undergoing mammography. *Am J Gastroenterol*. 2003;98(8):1875-1880.
30. Geiger TM, Miedema BW, Geana MV, Thaler K, Rangnekar NJ, Cameron GT. Improving rates for screening colonoscopy: analysis of the health information national trends survey (HINTS I) data. *Surg Endosc*. 2008;22(2):527-533.
31. van Roon AH, Hol L, Wilschut JA, et al. Advance notification letters increase adherence in colorectal cancer screening: a population-based randomized trial. *Prev Med*. 2011;52(6):448-451.
32. Marcus AC, Mason M, Wolfe P, et al. The efficacy of tailored print materials in promoting colorectal cancer screening: results from a randomized trial involving callers to the National Cancer Institute's Cancer Information Service. *J Health Commun*. 2005;10(suppl 1):83-104.
33. Courtney RJ, Paul CL, Sanson-Fisher RW, Carey ML, Macrae FA, Yoong SL. Community approaches to increasing colorectal screening uptake: a review of the methodological quality and strength of current evidence. *Cancer Forum*. 2012;36(1):27-35.
34. Powe BD, Ntekop E, Barron M. An intervention study to increase colorectal cancer knowledge and screening among community elders. *Public Health Nurs*. 2004;21(5):435-442.
35. Basch CE, Wolf RL, Brouse CH, et al. Telephone outreach to increase colorectal cancer screening in an urban minority population. *Am J Public Health*. 2006;96(12):2246-2253.
36. Harewood GC, Wiersema MJ, Melton LJ 3rd. A prospective, controlled assessment of factors influencing acceptance of screening colonoscopy. *Am J Gastroenterol*. 2002;97(12):3186-3194.
37. Wilf-Miron R, Peled R, Yaari E, Vainer A, Porath A, Kokia E. The association between socio-demographic characteristics and adherence to breast and colorectal cancer screening: analysis of large sub populations. *BMC Cancer*. 2011;11:376.
38. Bentler PM, Speckart G. Models of attitude-behavior relations. *Psychological Review*. 1979;86(5):452-464.
39. Bagozzi RP. Attitudes, intentions, and behavior: a test of some key hypotheses. *Journal of Personality and Social Psychology*. 1981;41(4):607-627.
40. Fredericks AJ, Dossett DL. Attitude-behavior relations: a comparison of the Fishbein-Ajzen and the Bentler-Speckart models. *Journal of Personality and Social Psychology*. 1983;45(3):501-512.
41. Bagozzi RP, Warshaw PR. Trying to consume. *Journal of Consumer Research*. 1990;17(2):127-140.
42. *SPSS Decision Trees* [software]. Version 17.0. Chicago, IL: SPSS Inc; 2007.
43. Perlch C, Provost F, Simonoff JS. Tree induction vs. logistic regression: a learning-curve analysis. *Journal of Machine Learning Research*. 2003;4:211-255.
44. Kansagara D, Englander H, Salanitro A, et al. Risk prediction models for hospital readmission: a systematic review. *JAMA*. 2011;306(15):1688-1698. ■

www.ajmc.com Full text and PDF
Web exclusive eAppendix

eAppendix

eAppendix Table 1: Segment 1 Description for Each Group

	Holdout Group (n = 400)	Intervention Group (n = 14,618)
Gender		
• Female	224 (56.00%)	7987 (54.64%)
• Male	176 (44.99%)	6631 (45.36%)
Age (Mean)	$\mu = 56.8$	$\mu = 57.3$
Standard deviation	$\sigma = 4.9$	$\sigma = 5.1$
Personal or family history of CRC		
• ICD-9 code V16.0, V18.5, V12.72	0 (0%)	25 (0.17%)
• Not coded	400 (100.0%)	14,770 (99.83%)
Physical exam in the last 2 years		
• Yes	30 (7.5%)	1381 (9.45%)
• No	370 (92.5%)	13,237 (90.55%)
Undiagnosed abdominal pain		
* ICD-9 code 789.0X	10 (2.50%)	422 (2.89%)
* Not coded	390 (97.50%)	14,196 (97.11%)
Breast cancer screening		
• Ineligible	1 (0.45%)	162 (2.03%)
• Not completed	209 (93.72%)	7295 (91.31%)
• Completed	13 (5.83%)	532 (6.66%)
Emergency department visit		
• No visit	26 (6.5%)	1,458 (9.97%)
• At least 1 visit in 2 years	374 (93.50%)	13,160 (90.03%)

CRC indicates colorectal cancer; ICD-9, International Classification of Diseases, Ninth Revision, Clinical Modification.

eAppendix Table 2: Segment 2 Description for Each Group

	Holdout Group (n = 313)	Intervention Group (n = 11,164)
Gender		
• Female	158 (50.48%)	5,623 (54.64%)
• Male	155 (49.52%)	5,541 (49.58%)
Age (mean)	$\mu = 58.8$	$\mu = 58.8$
Standard deviation	$\sigma = 4.9$	$\sigma = 5.2$
Personal or family history of CRC		
• ICD-9 code V16.0, V18.5, V12.72	0 (0%)	9 (.08%)
• Not coded	313 (100.0%)	11,155 (99.92%)
Physical exam in the last 2 years		
• Yes	82 (26.20%)	3062 (27.43%)
• No	231 (73.80%)	8102 (72.57%)
Undiagnosed abdominal pain		
* ICD-9 code 789.0X	22 (7.03%)	776 (6.95%)
* Not coded	291 (92.97%)	10,388 (93.05%)
Breast cancer screening		
• Ineligible	32 (20.13%)	1364 (24.24%)
• Not completed	64 (40.25%)	2194 (39.00%)
• Completed	63 (39.62%)	2068 (36.76%)
Emergency department visit		
• No visit	284 (90.73%)	10,721 (89.81%)
• At least 1 visit in 2 years	29 (9.27%)	1216 (10.19%)

CRC indicates colorectal cancer; ICD-9, *International Classification of Diseases, Ninth Revision, Clinical Modification*.

eAppendix Table 3. Segment 3 Description for Each Group

	Holdout Group (n = 752)	Intervention Group (n = 11,268)
Gender		
• Female	377 (50.13%)	4598 (40.80%)
• Male	375 (49.87%)	6670 (59.20%)
Age (mean)	$\mu = 58.4$	$\mu = 58.7$
Standard deviation	$\sigma = 5.84$	$\sigma = 5.5$
Personal or family history of CRC		
• ICD-9 code V16.0, V18.5, V12.72	0 (0%)	4 (.04%)
• Not coded	752 (100.0%)	11,264 (99.96%)
Physical exam in the last 2 years		
• Yes	291 (38.70%)	3336 (29.61%)
• No	461 (61.30%)	7932 (70.39%)
Undiagnosed abdominal pain		
* ICD-9 code 789.0X	82 (10.90%)	1275 (11.32%)
* Not coded	670 (89.10%)	9993 (88.68%)
Breast cancer screening		
• Ineligible	34 (9.02%)	371 (12.53%)
• Not completed	31 (8.22%)	522 (11.35%)
• Completed	312 (82.76%)	3705(80.58%)
Emergency department visit		
• No visit	569 (75.66%)	8473 (75.20%)
• At least 1 visit in 2 years	183 (24.34%)	2795 (24.80%)

CRC indicates colorectal cancer; ICD-9, *International Classification of Diseases, Ninth Revision, Clinical Modification*.

eAppendix Table 4. Segment 4 Description for Each Group

	Holdout Group (n = 472)	Intervention Group (n = 5838)
Gender		
• Female	293 (62.08%)	3642 (62.38%)
• Male	179 (37.92%)	2196 (37.62%)
Age (mean)	$\mu = 58.2$	$\mu = 58.3$
Standard deviation	$\sigma = 5.25$	$\sigma = 5.32$
Personal or Family History of CRC		
• ICD-9 code V16.0, V18.5, V12.72	0 (0%)	0 (0%)
• Not coded	472 (100.0%)	5838 (100.00%)
Physical exam in the last 2 years		
• Yes	361 (76.48%)	4232 (72.49%)
• No	111 (23.52%)	1606 (27.51%)
Undiagnosed abdominal pain		
* ICD-9 code 789.0X	141 (29.87%)	1573 (26.94%)
* Not coded	331 (70.13%)	4265 (73.06%)
Breast cancer screening		
• Ineligible	7 (2.39%)	75 (2.06%)
• Not completed	5 (1.71%)	54 (1.48%)
• Completed	281 (95.90%)	3512(96.46%)
Emergency department visit		
• No visit	324 (68.64%)	4265 (73.06%)
• At least 1 visit in 2 years	148 (31.36%)	1573 (26.94%)

CRC indicates colorectal cancer; ICD-9, *International Classification of Diseases, Ninth Revision, Clinical Modification*.

eAppendix Table 5: Segment 5

	Holdout Group (n = 63)	Intervention Group (n = 915)
Gender		
• Female	54 (85.71%)	732 (80.00%)
• Male	9 (14.29%)	183 (20.00%)
Age (mean)	$\mu = 58.2$	$\mu = 58.6$
Standard deviation	$\sigma = 5.27$	$\sigma = 5.34$
Personal or family history of CRC		
• ICD-9 code V16.0, V18.5, V12.72	16 (25.40%)	260 (28.42%)
• Not coded	47 (74.60%)	655 (71.58%)
Physical exam in the last 2 years		
• Yes	53 (84.13%)	740 (80.87%)
• No	10 (15.87%)	175 (19.13%)
Undiagnosed abdominal pain		
* ICD-9 code 789.0X	26 (41.27%)	392 (42.84%)
* Not coded	37 (58.73%)	523 (57.16%)
Breast cancer screening		
• Ineligible	31 (57.41%)	375 (51.23%)
• Not completed	1 (1.85%)	23 (3.14%)
• Completed	22 (40.74%)	334 (45.63%)
Emergency department visit		
• No visit	45 (71.43%)	595 (65.03%)
• At least 1 visit in 2 years	18 (28.57%)	320 (34.97%)

CRC indicates colorectal cancer; ICD-9, International Classification of Diseases, Ninth Revision, Clinical Modification.