

Interventions to Promote Repeat Breast Cancer Screening With Mammography: A Systematic Review and Meta-Analysis

Sally W. Vernon, Amy McQueen, Jasmin A. Tiro, Deborah J. del Junco

Manuscript received May 20, 2009; revised May 19, 2010; accepted May 20, 2010.

Correspondence to: Sally W. Vernon, PhD, Center for Health Promotion and Prevention Research, Division of Health Promotion and Behavioral Sciences, University of Texas School of Public Health, 7000 Fannin, Ste 2560, Houston, TX 77030 (e-mail: sally.w.vernon@uth.tmc.edu).

Background Various interventions to promote repeat use of mammography have been evaluated, but the efficacy of such interventions is not well understood.

Methods We searched electronic databases through August 15, 2009, and extracted data to calculate unadjusted effect estimates (odds ratios [ORs] and 95% confidence intervals [CIs]). Eligible studies were those that reported estimates of repeat screening for intervention and control groups. We tested homogeneity and computed summary odds ratios. To explore possible causes of heterogeneity, we performed stratified analyses, examined meta-regression models for 15 a priori explanatory variables, and conducted influence analyses. We used funnel plots and asymmetry tests to assess publication bias. Statistical tests were two-sided.

Results The 25 eligible studies (27 effect estimates) were statistically significantly heterogeneous ($Q = 69.5$, $I^2 = 63\%$, $P < .001$). Although there were homogeneous subgroups in some categories of the 15 explanatory variables, heterogeneity persisted after stratification. For all but one explanatory variable, subgroup summary odds ratios were similar with overlapping confidence intervals. The summary odds ratio for the eight heterogeneous reminder-only studies was the largest observed (OR = 1.79, 95% CI = 1.41 to 2.29) and was statistically significantly greater than the summary odds ratio ($P_{\text{diff}} = .008$) for the homogeneous group of 17 studies that used the more intensive strategies of education/motivation or counseling (OR = 1.27, 95% CI = 1.17 to 1.37). However, reminder-only studies remained statistically significantly heterogeneous, whereas the studies classified as education/motivation or counseling were homogeneous. Similarly, in meta-regression modeling, the only statistically significant predictor of the intervention effect size was intervention strategy (reminder-only vs the other two combined as the referent). Publication bias was not apparent.

Conclusions The observed heterogeneity precludes a summary effect estimate. We also cannot conclude that reminder-only intervention strategies are more effective than alternate strategies. Additional studies are needed to identify methods or strategies that could increase repeat mammography.

J Natl Cancer Inst 2010;102:1023–1039

Breast cancer is the second leading cause of cancer deaths in women in the United States (1). Regular screening with mammography has been shown to reduce mortality from breast cancer in women aged 50–74 years by approximately 23% (2). To maximize the population benefit related to mortality reduction, the US Preventive Services Task Force recommended in 2002 that women aged 40 years and older be screened with mammography every 1–2 years. Although the Task Force has raised the minimum age for biennial screening to 50 years of age, they still suggest that younger women discuss mammography with their doctors to make an informed decision based on their family histories, personal values, and general health (<http://www.ahrq.gov/clinic/uspstf09/breastcancer/brcanrs.htm>) (3). Breast cancer screening with mammography increased substantially since 1990. Surveillance data from the National Health Interview Survey show that the prevalence of self-

reported recent use (within the past 2 years) in women 40 years of age and older increased from 30% in 1987 to 70% in 2000 (4); however, data from the 2005 National Health Interview Survey show a decline to 66% (5). The prevalence of regular or repeat mammography use, that is, consecutive, on-schedule mammograms, is lower compared with recent use, that is, one mammogram within the past 2 years. A review of 37 regional studies of repeat mammography conducted through 2001 found that the overall weighted average prevalence was 46.1% (95% confidence interval [CI] = 39.4 to 52.8) (6). Summary estimates also showed that repeat use increased from 26.5% (95% CI = 12.9 to 40.0) in studies conducted before 1991 to 53.2% (95% CI = 44.7 to 61.8) in those conducted between 1995 and 2001 (6).

Systematic reviews and meta-analyses have shown that many intervention strategies are effective at motivating women to have

CONTEXTS AND CAVEATS

Prior knowledge

Regular mammography screening has been shown to reduce mortality from breast cancer in women aged 50–74 years, but it is not known whether interventions to promote regular screening, for example, reminders, educational outreach, and counseling, are effective.

Study design

The effectiveness of various intervention strategies was examined in a systematic review and meta-analysis of studies that reported estimates of repeat screening for intervention and control groups.

Contribution

The 25 analyzed studies were heterogeneous overall, as were most subgroups in a stratified analysis. The intervention effect of a group of reminder-only studies was statistically significantly greater than that of a group of studies that used more intensive strategies such as education/motivation or counseling.

Implications

Heterogeneity prevents firm conclusions about the effectiveness of more intensive vs less intensive strategies. More studies with consistent designs and well-defined intervention categories are needed.

Limitations

Not all studies compared participants with dropouts or reported differential attrition by study group, which may result in overestimated effect sizes. Most of the studies were conducted more than 10 years ago, and most were of non-Hispanic white women, so the results may not be applicable to the present day or to other ethnic groups.

From the Editors

one mammogram during the study period (7–19). Most meta-analyses consistently show that minimal interventions such as reminders directed at patients (7,8,10–12,14,17,19) or providers (13–15,18) delivered through a variety of communication channels are effective in increasing one-time mammography screening compared with a no-intervention control group. Systematic reviews and meta-analyses also demonstrate that more intensive patient-directed interventions, including those using multiple strategies (eg, letter plus telephone call, letter plus voucher), those tailored to an individual's beliefs or characteristics (eg, a personalized message that addresses a woman's concerns such as fear of finding cancer), and those based on health behavior theories (a list of theories commonly used in health promotion research is available at <http://dccps.cancer.gov/brp/constructs>) reported larger intervention effects compared with a no-intervention control group or with minimal interventions, such as mail or telephone reminders (7,11,12,16,19). Reducing barriers to access such as cost or transportation also is associated with increased mammography use (8,17).

In contrast to interventions to promote one-time screening, the efficacy of interventions designed to promote regular mammography screening is not well understood (20). To our knowledge, there is no systematic review or meta-analysis of interventions that promote repeat mammography screening, perhaps because fewer

interventions have reported repeat mammography outcomes compared with one-time use. Developing approaches to encourage women to maintain a regular schedule of mammography screening is needed if we are to realize a reduction in breast cancer mortality.

We conducted a systematic review of repeat mammography intervention studies, in which we evaluated sampling, methods, and intervention characteristics. We examined consistency of effect across studies and assessed completeness of reporting on selected study characteristics related to internal and external validity. We also identified gaps in the literature and make recommendations for further research.

Methods

Search Strategy

We conducted electronic database searches in consultation with a medical librarian who was trained in systematic review literature searches to identify articles reporting the effects of behavioral interventions on repeat mammography. We concluded the database searches on August 15, 2009. First, we searched MEDLINE (OVID) from 1966. Then, we adapted the search for CINAHL (OVID) from 1982, PsycInfo (OVID) from 1967, and Academic Search Premier (EBSCO) from 1990. We repeated the search terms used in the Clark et al. review (6), which paired “mammogra\$” and seven keywords: “regular,” “repeat,” “adherence,” “compliance,” “annual,” “rescreen,” and “maintenance” (“maint\$”); we added “biennial,” “on schedule,” and “guideline\$.” We used several search terms to identify controlled trials or interventions, as recommended by the Cochrane Collaboration (21). We combined search results from the two previous steps and then limited the results to human studies published in English that were not editorials, commentaries, letters to the editor, reviews, or meta-analyses. Using Medical Subject Headings terms (<http://scientific.thomson.com/support/faq/wok3new/medline/#MeSH>), we further excluded studies that focused on diagnostic techniques and procedures, biopsy, and drug therapy. Reference lists from eligible articles and observational studies of repeat mammography were hand searched for additional reports. We also hand searched the Cochrane databases (<http://www.thecochranelibrary.com/view/0/index.html>) for eligible studies, as well as literature reviews of mammography interventions.

Selection Criteria

To be included, studies had to report an estimate of repeat mammography use for at least one intervention group and one concurrent comparison group. We used the Clark et al. (6) definition of repeat screening, that is, at least two consecutive, on-schedule mammograms during a given period (approximately 1–2 years apart); a certain number of mammograms during a given period (at least two within the past 5 years); or at least two mammograms on an age-appropriate schedule (eg, biennially for women in their forties). Like Clark et al. (6), we excluded studies defining repeat mammography as more than one lifetime mammogram without regard to period, receipt of a single mammogram with an intention to obtain future screening, or behaviors of health-care providers such as a recommendation to get screened. The intervention had to include

women at average risk for breast cancer, but it also could address other health-related behaviors in addition to mammography, such as other cancer screening behaviors, smoking, or physical activity.

Two authors independently reviewed titles and abstracts to identify relevant articles according to the inclusion and exclusion criteria. Full-text articles were reviewed for eligibility when more information was needed. A hierarchical categorization scheme was applied to each study to determine eligibility. Studies were first classified as including an intervention or no intervention. Intervention studies were considered for inclusion and further classified as having repeat mammography as an outcome. Disagreements about study eligibility were discussed with all coauthors until consensus was reached.

Data Extraction to Describe Sampling, Methods, and Intervention Characteristics of Eligible Studies

We used a standardized data extraction form to record descriptive information from eligible studies. In studies of women with diverse mammography histories, we included only the subgroup of women who were eligible to complete at least one repeat mammogram. Specifically, we identified two types of study designs used in intervention studies of repeat mammography. In the most common type of design, study participants with a recent (ie, not overdue) mammogram before the intervention were followed long enough to receive one postintervention mammogram on schedule (hereafter referred to as design 1). In the other design, study subjects were followed long enough to receive two postintervention mammograms on schedule regardless of mammography history (hereafter referred to as design 2).

We extracted information on the following study characteristics: year(s) the study was conducted, age range of participants, race/ethnicity (percent white was used rather than a racial or ethnic breakdown because most of the studies reporting racial/ethnic composition of the study population had studied white women), study setting (health care or community), percent of the sample with a recent mammogram at study baseline, screening interval used to measure adherence (1 year, 2 years, or age dependent [biennially for women in their forties and annually or biennially for women in their fifties]), study design type (1 or 2), data source for mammography status (medical records, administrative or program data, or self-report), intervention strategy, mode of intervention delivery (mail only, telephone only, mail plus telephone, mail plus in person, or community education plus other modes), number of study groups, type of control group (no contact, survey only, or active [alternative intervention of equal or lower intensity]), theoretic framework [eg, health belief model (22)], and theoretic constructs (eg, barriers such as cost and stage of change or readiness to be screened [<http://dccps.cancer.gov/brp/constructs>]).

For studies that included more than one intervention group, we abstracted information on the intervention strategy of greatest intensity (eg, personalized vs generic messages or multiple vs single strategies) or based on the author's hypothesis. For example, Finney and Iannotti (23) tested messages framed in terms of what would be gained or lost by getting a mammogram and hypothesized that messages emphasizing what would be lost by not being screened would be more effective than messages emphasizing what would be gained. We also abstracted information on delivery mode

and on theoretic frameworks and constructs only for the group that received the most intensive intervention. Where relevant, we identified the theoretic constructs used to deliver tailored messages. For example, Clark et al. (6) created letters using each woman's interview responses to questions about her stage of change or readiness to get a mammogram and her perceived benefits of and barriers to getting screened.

A variety of approaches have been used to describe and classify intervention strategies, and no one approach is considered the gold standard (24). Because there is no agreed-upon classification, we classified intervention strategies in three ways. Our primary classification was informed by health behavior theory (<http://dccps.cancer.gov/brp/constructs>) and consisted of three categories: reminder, education/motivation, and counseling. We based this classification on the work of Kreuter et al. (25,26) who describe intervention strategies and messages in terms of their personal relevance to the recipient. In general, generic messages such as mailed reminders are considered to be the least intensive, whereas communication based on an assessment of a person's beliefs and attitudes and delivered through interpersonal communication channels is considered to be the most intensive. Reminders, whether generic or personalized, consisted of minimal print or telephone messages that served as a cue to action or prompt by letting women know that they were due for screening. Reminders could contain minimal information such as a statement that a woman was due for screening and should call to schedule an appointment, or they could contain brief motivational messages based on health behavior theory, along with the reminder. Educational/motivational strategies consisted of print messages to increase knowledge, facilitate attitude change, and motivate women to be screened. Messages may or may not be personalized on the basis of information obtained from personal assessments (eg, surveys or interviews) or from other data sources such as medical records. Counseling strategies, typically delivered over the telephone or in person, are considered to be the most personalized form of communication because they engage women in a dialogue in an attempt to change attitudes, address barriers such as perceived risk of developing breast cancer or fear of pain from the mammogram, and motivate women to be screened. If an intervention used multiple strategies, we classified it based on the most intensive strategy.

Our second approach to classifying intervention strategies was to contrast the subset of studies that explicitly stated that they used barriers-specific telephone counseling with those that did not. Barriers-specific telephone counseling is an intervention approach used by the National Cancer Institute's Breast Cancer Screening Consortium (27), in which a counselor uses a standardized protocol to identify a person's barriers to performing a health behavior and provides information to address and overcome the barriers. In the third approach, we classified intervention strategies based on whether the study used a single intervention strategy (eg, reminder only) or multiple strategies (eg, reminder plus another strategy). We also classified intervention delivery in two ways: by delivery channel (eg, mail only, telephone only) and by whether a single mode or multiple modes were used.

Eligible studies were assigned to one coauthor (S. W. Vernon, A. McQueen, J. A. Tiro) for review. To assess reliability, one coauthor

(D. J. del Junco) reviewed a randomly selected 10% sample of the studies; disagreements were discussed by all authors until consensus was reached. Data to calculate effect size estimates were independently extracted by two coauthors, and disagreements were discussed by all authors until consensus was reached. When available, we reviewed other published reports of the same study; however, in only one instance, did we gain additional relevant information. In that case, companion articles reported on different aspects of a single study, and we treated these reports as one study (28,29). We did not contact authors to obtain unpublished data or missing information because that source is not readily available to readers and because we wanted to evaluate the extent to which selected study characteristics were reported in the published literature.

Statistical Analysis of Intervention Effects and Potential Explanatory Variables

Every study produced one or more comparisons of an intervention group with a control group. When there was more than one intervention condition, we contrasted the most intensive intervention condition with the control condition. Odds ratio (OR) effect sizes and 95% confidence intervals were calculated using cell frequencies or proportions from 2×2 contingency tables (30). We recalculated unadjusted odds ratios and confidence intervals for all studies; however, for two studies (31,32), insufficient data were available, and we used adjusted estimates reported by the authors. Statistical tests were two-sided.

Tests for measuring heterogeneity were conducted using the Q statistic with a $P < .05$ criterion and an I^2 statistic with a cutoff greater than or equal to 50% to indicate substantial heterogeneity (33). A statistically significant Q indicates a heterogeneous distribution of study effect sizes, which may then warrant additional subgroup analyses (30). The I^2 statistic describes the percentage of the variability in effect estimates because of heterogeneity rather than sampling error or chance alone (33). Heterogeneity tests were not performed when there were fewer than five observations in a category of a variable. Heterogeneity tests were performed using STATA 10.0 (34).

Variance-weighted summary effect sizes were computed. We used a fixed-effects model to summarize homogeneous distributions. For heterogeneous distributions, we report summarizations based on random-effects models (35). We also performed random-effects meta-regression analyses using STATA on potential explanatory variables of intervention effects (35,36). Meta-regressions were performed with each variable univariately, with a forward variable selection procedure that included variables with a univariate P less than .25 and eliminated variables with a multivariable-adjusted P greater than .05.

For the heterogeneity and meta-regression analyses, we based our a priori choice of potential explanatory variables on prior systematic reviews of studies that measured completion of one mammogram during the study period (8,16) and on factors associated with repeat screening examined by Clark et al. (6). We examined 15 covariates (categorization of these variables is described earlier): age, study setting, screening interval, study design type, data source for repeat mammography outcome, intervention strategy (classified in three ways, as described earlier), mode of intervention delivery, number of delivery modes, control group type, use of a

theoretic framework, two theoretic constructs (barriers and stage of change), and use of tailoring (personalizing the message).

Most studies used multiple theoretic frameworks, and there were not enough studies using a given framework to form reliable groupings. For example, only three studies used the transtheoretical model (37) alone and only one study used the health belief model (22) alone, the two most frequently cited models. Therefore, we classified studies in terms of whether a theoretic framework was used (yes or no). Likewise, with only a few exceptions, very few studies measured the same theoretic constructs. Two exceptions were the barriers construct from the health belief model (22), which was measured in 13 studies, and the stage of change construct from the transtheoretical model (37), which was measured in nine studies. Those were the only two theoretic constructs included in our analyses. We also identified studies that tailored messages on the basis of one or more constructs, for example, knowledge and stage of change. We created a variable called use of tailoring and classified studies as yes or no on that variable.

To examine the contribution of individual studies to the overall summary effect estimate, we conducted an influence analysis (omitting one study at a time) (38). The influence analysis produces a graph enabling the assessment of the influence of one study on the overall meta-analysis summary odds ratio estimate by visually comparing summary effect estimates after the removal of each study's effect estimate on successive turns. To assess the potential for publication bias, we performed funnel plot asymmetry tests (39,40). The Begg test (39) is directly analogous to a visual assessment of funnel plot symmetry (ie, the dispersion of all point estimates from all studies on a graph to form a symmetrical funnel shape), and it tests whether the Begg rank correlation between effect size and its SE is zero. Pseudo confidence intervals are the points connected by the diagonal lines forming the "funnel" on the funnel plot; they are the expected 95% confidence intervals for a given SE (depicted as increasing along the x -axis). The Egger test (40) is a regression of the standardized effect size (eg, log OR/SE of log OR) against its precision (eg, $1/SE$ of log OR). If the intercept of the Egger regression line differs statistically significantly from zero, publication bias may be present.

Completeness of Reporting on Selected Aspects of Internal and External Validity

The importance of systematically assessing aspects of internal validity in health promotion trials has been recognized for some time (41). One aspect of internal validity that has recently been emphasized is the design and analysis of group- or cluster-randomized trials in cancer prevention and control (42). Recent attention also has been directed to the importance of assessing external validity in health promotion research, and several frameworks for assessing it have been proposed (43,44). We assessed studies for completeness of reporting on aspects of internal and external validity related to representativeness of the study population and to design and analysis issues. To assess representativeness, we recorded whether or not authors provided information on the response rate at baseline, comparison of respondents and nonrespondents at baseline, equivalence of study groups at baseline, response rate at follow-up, comparison of the final sample to dropouts, and whether there was differential attrition by study group. To assess design and analysis

issues, we recorded whether the authors provided information on how the sample size was determined (ie, a priori statistical power analysis), whether an intent-to-treat analysis was performed, the unit of randomization, and, for group-randomized trials, whether the outcome analysis was adjusted for nested data.

Results

We identified 319 unique articles that matched our search criteria. Of the 319 articles reviewed for eligibility, 165 studies were excluded because they were not intervention studies (Figure 1). Of the 153 intervention studies, 39 reported repeat mammography outcomes. Thirteen of the 39 were excluded because the study design (45–49), analysis or reporting (50–55), or study sample (56,57) could not be directly compared with the other repeat mammography studies. In one study that stratified results by family history of breast cancer (23), we extracted data only for the group of women without a family history because our focus was on women at average risk. We reached consensus that 25 studies (26 articles) were eligible for review (Figure 1).

Description of Eligible Studies

Study population and setting. Most studies were conducted in the mid-to-late 1990s or later; four did not report when the study was conducted (Table 1). Nine studies included women with a minimum age of 40 years, 14 with a minimum age of 50–52 years, and two with a minimum age of 65 years. Fifteen studies imposed an upper age limit, whereas the rest did not.

Of the 19 studies reporting racial/ethnic composition of the study population, most studied non-Hispanic white women. Seventeen of the 25 studies recruited women through health-care settings (health maintenance organizations, primary care practices, outpatient clinics, or mammography facilities). Of the eight studies

that recruited community residents, five used population-based sampling strategies (29,58–61), two recruited through the state’s Breast and Cervical Cancer Early Detection Program (62,63), and one through churches (64).

Most studies recruited women with diverse mammography histories, but 11 (23,59,62,63,65–71) recruited only women who had an up-to-date mammogram and would become due for another one during the study period. In studies of women with diverse mammography histories, the percentage of women with an up-to-date mammogram at study baseline (and therefore eligible to be included in our effect size estimates) ranged from 4% (31) to 72% [(32), Table 1].

Measurement of repeat mammography. Mammography adherence was assessed as annual use in 17 studies, biennial use in six studies, and age dependent (ie, biennial for women <50 years of age and annual for women ≥50 years of age) in two studies (Table 1). Eighteen studies used design 1, that is, they measured one pre- and one postintervention mammogram. Five studies used design 2, that is, they measured two on-schedule postintervention mammograms, and two studies reported data in a way that permitted calculation of repeat mammography estimates for both design types. Thus, a total of 27 individual effect size estimates were available for analysis. Thirteen of the 25 studies used some type of objective record data to measure the outcome of mammography completion, nine used self-report, and three used a combination of records and self-report (Table 1).

Description of the interventions. All but two studies were randomized controlled trials. Eaker et al. (60) selected eight counties and designated four as intervention and four as control. Quinley et al. (71) identified mammography facilities that used reminders and compared them with facilities that did not. Eight studies used only patient reminders as the primary intervention

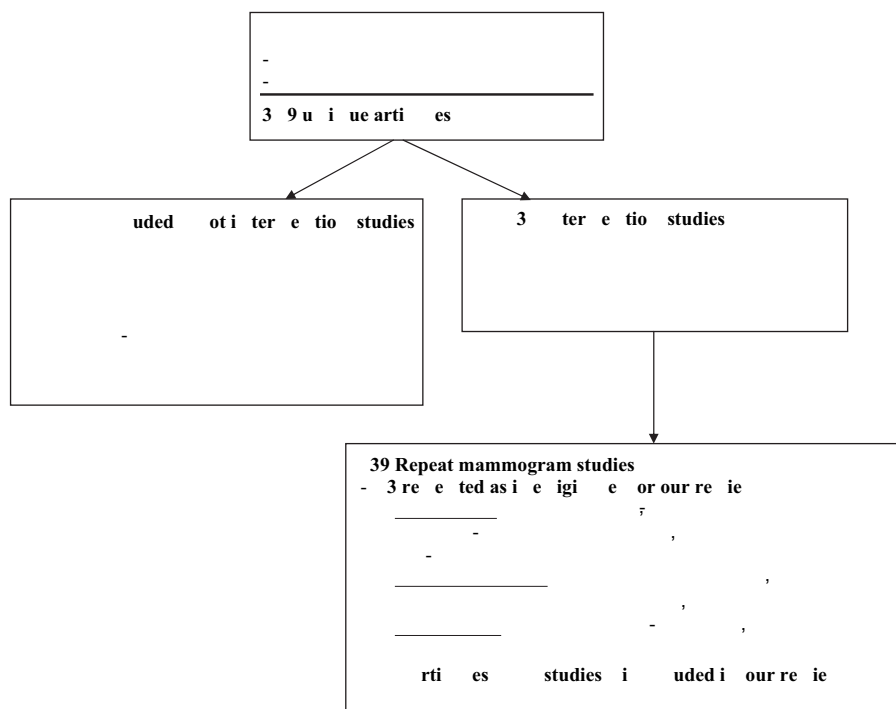


Figure 1. Literature search results and review process of publications by exclusionary criteria. The category “other” under excluded nonintervention studies includes measurement, risk assessment models, perceptions and knowledge, breast self-examinations, abnormal mammography results and postmammogram psychological reactions (eg, anxiety, worry), patient satisfaction, and genetic screening.

Table 1. Characteristics of 25 studies of interventions to promote repeat mammography screening*

First author (reference)	Study population and setting			Measurement of repeat mammography			Intervention description						
	Study years	Age range, y	% Non-Hispanic white	Study setting	% Up-to-date at baseline†	Screening interval, y	Design†	Data source for screening outcome	Strategy	Delivery mode	No. of study groups	Control group	Theoretic framework
Andersen (61)	1994–1997	50–80	97	Community	~50	2	1	SR	Education/motivation + counseling	Community education + telephone	4	Survey only	HBM, conflict model of decision making
Barr (65)	1994–1996	50–75	NR	Health care	100	2	1	AD and MR	Reminder	Telephone	3	Active	HBM, TRA
Bodiya (66)	NR	≥50	NR	Health care	100	1	1	AD	Reminder	Mail + phone	3	No contact	None
Clark (72)	1993–1996	50–74	95	Health care	54	1	1 and 2	MR	Education/motivation	Mail	3	Survey only	TTM
Costanza (75)	1994–1997	50–80	97	Health care	~36	2	1 and 2	SR	Reminder + counseling	Mail + telephone	3	Active	HBM, TTM, conflict model of decision making
Crane (58)	1994–1996	≥50	79	Community	63	2	1	SR	Education/motivation + counseling	Mail + telephone	3	Survey only	TTM
DeFrank (70)	2004–2006	40–75	88	Health care	100	1	1	AD and SR	Reminder + education/motivation	Mail	3	Active	HBM
Drossaert (59)	NR	50–69	NR	Community	100	2	1	Program records	Education/motivation	Mail	3	Active	Attitude-social influence-efficacy model
Duan (64)	1996–1998	50–80	53	Community	~65	1	1	SR	Counseling	Telephone	2	Survey only	HBM
Eaker (60)	1993–1996	≥40	NR	Community	51	1: ≥50 y 2: 40–49 y	1	SR	Reminder + education/motivation	Community education + mail	2	Survey only	None
Finney (23)	1999–2000	40–69	93	Health care	100	1	1	MR	Reminder	Mail	3	Active	Prospect theory
Goel (62)	2004	40–64	23	Community	100	1	1	Program records	Reminder	Mail + telephone	2	Active	None
Lipkus (32)	1994–1997	≥50	82	Health care	72	1	2	SR	Reminder + counseling	Telephone	3	Active	HBM, TTM, conflict model of decision making
Margolis (73)	1992–1995	≥40	62	Health care	44	1: ≥50 y 2: 40–49 y	1	AD and SR	Lay advisor/navigator	Mail + in person	2	Survey only	None
Mayer (67)	1995–1998	50–74	83	Health care	100	1	1	MR	Reminder	Mail	3	Survey only	TRA, operant learning theory

(Table continues)

Table 1 (continued).

First author (reference)	Study population and setting			Measurement of repeat mammography			Intervention description						
	Study years	Age range, y	% Non-Hispanic white	Study setting	% Up-to-date at baseline†	Screening interval, y	Design‡	Data source for screening outcome	Strategy	Delivery mode	No. of study groups	Control group	Theoretic framework
Messina (76)	1995–1998	50–80	97	Health care	32	2	2	SR	Education/motivation + counseling	Mail + telephone	2	Survey only	TTM
Michielutte (74)	1999–2002	≥65	86–93	Health care	31	1	1	MR	Education/motivation + counseling	Mail + telephone	2	Active	HBM, TTM
Partin (63)	1998–2000	40–63	89	Community	100	1	1	Program records	Reminder + education/motivation	Mail	3	Survey only	None
Quinley (71)	1999–2000	≥65	90	Health care	100	1	1	AD	Reminder	Mail	2	No contact	None
Rakowski (68)	1996–1998	50–74	88	Health care	100	1	1	AD	Reminder + education/motivation + counseling	Mail + telephone	4	Active	HBM, TTM, conflict model of decision making
Rimer (77)	1997–2000	40–44, 50–54	82	Health care	67	1	2	SR	Reminder + education/motivation + counseling	Mail + telephone	3	Active	ELM, TTM, PAM
Schapira (69)	NR	40–70	NR	Health care	100	1	1	MR and SR	Reminder	Mail + in person	4	Active	None
Simon (31)	1992–1993	≥40	NR	Health care	4	1	1	MR	Reminder	Mail	3	Active	None
Skinner (78)	NR	≥51	48	Health care	49	1	2	MR	Education/motivation + counseling	Mail + telephone	4	Survey only	HBM, TTM
Vernon and del Junco (29, 28)	2000–2004	≥52	85	Community	46	1	2	SR	Education/motivation	Mail	3	Survey only	HBM, TTM, TRA/TPB, SCT

* Of the 19 studies reporting racial/ethnic composition, most studied non-Hispanic white women. AD = administrative data; ELM = elaboration likelihood model; HBM = health belief model; MR = medical record; NR = not reported; PAM = precaution adoption process model; SCT = Social Cognitive Theory; SR = self-report; TRA/TPB = theory of reasoned action/planned behavior; TTM = transtheoretical model.

† In studies where 100% of the women had an up-to-date mammogram at baseline, the sample was selected to include only women who were not overdue for their next mammogram but would become due during the study period. In some instances, estimates had to be averaged across groups and so these are approximate estimates.

‡ Design 1 studies measured one pre- and one postintervention mammogram; design 2 studies measured two on-schedule postintervention mammograms.

strategy (23,31,62,65–67,69,71), six used educational or motivational strategies (29,59,60,63,70,72), and 11 (32,58,61,64,68,73–78) involved in-person or telephone counseling (Table 1). The one study (73) that used lay health advisors or navigators was grouped with the counseling interventions for subsequent analysis because their role was to educate women about breast cancer and screening and to assist women in scheduling and completing a mammogram. Variability existed within intervention strategy. For example, reminders to complete a mammogram varied by source (ie, facility or personal physician), tailoring or framing of the message, and timing. About half of the 25 studies used multiple intervention strategies, and eight (32,58,61,68,74–77) of the 11 counseling studies used some variation of barriers-specific telephone counseling. Intervention delivery mode also varied across studies (Table 1); nine involved only mail, three used only telephone, nine used both mail and telephone, two involved mail and in-person contact, and two used a variety of community education strategies in addition to mail or telephone.

The number of study groups varied, with some studies using multiple intervention and comparison groups (Table 1). Eleven studies used a survey-only control group, whereas 12 had an active control or comparison intervention group. Only two studies included a no-contact control group.

Most studies (17 of 25) used a theoretic framework to guide the study intervention, and most studies that used a framework used more than one (Table 1). The most commonly used frameworks were the transtheoretical model (37) and the health belief model (22). The constructs most frequently used in interventions were barriers from the health belief model and stage of change from the transtheoretical model. Twelve studies tailored intervention messages on one or more theoretic constructs (29,32,58,61,64,68,72,74–78). Of the eight studies that did not explicitly identify a theoretic framework (31,60,62,63,66,69,71,73), all but one (73) used patient reminders as at least one of the intervention strategies, an approach consistent with the cue to action construct from the health belief model.

Effect Estimates for Interventions to Increase Repeat Mammography

Across all 25 eligible studies (27 effect estimates), the test for heterogeneity was statistically significant ($Q = 69.5$, $I^2 = 63\%$, $P < .001$; Table 2 and Figure 2). When subgroups of studies were classified under each of our 15 a priori categorical covariates, nine homogeneous subgroups were identified but only for certain categories of a covariate (Table 2): community study setting, design 2 (ie, two postintervention mammograms), self-report data for mammography completion, the three ways of classifying intervention strategies (use of education/motivation or counseling; use of barriers-specific telephone counseling; and use of multiple intervention strategies), use of the stage of change construct in the intervention, and use of tailoring. Statistically significant heterogeneity remained in all categories of the other seven covariates. With one exception, confidence intervals overlapped when comparing odds ratios across categories within each covariate, indicating similar subgroup effect sizes (ORs). The exception was the reminder-only intervention strategy (OR = 1.79, 95% CI = 1.41 to 2.29, $P < .001$) compared with the education/motivation

strategy (OR = 1.25, 95% CI = 1.14 to 1.38, $P = .868$) (Table 2). The summary odds ratio for the eight heterogeneous reminder-only studies was the largest observed (OR = 1.79, 95% CI = 1.41 to 2.29) and was statistically significantly greater than the summary odds ratio ($P_{diff} = .008$) for the homogeneous group of 17 studies that used the more intensive strategies of education/motivation or counseling (OR = 1.27, 95% CI = 1.17 to 1.37) (Table 2).

Variables with more than two categories were dichotomized for the meta-regression. For screening interval, the category age dependent, representing two studies (60,73), was combined with the category 1 year because the majority of women were aged 50 years or older, and 1 year is the commonly recommended interval. For data source, the three studies (69,70,73) that used a combination of self-report and medical records were combined with studies using medical records. For control group type, the two studies (66,71) that used a no-contact control group were combined with the survey-only category. For intervention strategy, the categories education/motivation and counseling were combined because their odds ratios were the same. Delivery mode was dichotomized as mail only vs other modes. In meta-regression modeling that included all 27 estimates, the only statistically significant predictor of the magnitude of the odds ratios was the intervention strategy of reminder only vs education/motivation and counseling combined as the referent (OR = 1.35, 95% CI = 1.08 to 1.68, $P = .011$). However, the overall I^2 value for this model remained 60.6%, indicating substantial residual heterogeneity and confirming the results of the heterogeneity analyses in Table 2 for the studies using reminder-only interventions.

In the influence analysis that included all 27 estimates, omitting the Mayer et al. (67) reminder-only study had the most pronounced effect (I^2 changed from 62.2% to 57.7%) but only slightly decreased the summary odds ratio estimate (from 1.39 to 1.35). In a separate influence analysis restricted to the eight reminder-only intervention studies that remained heterogeneous (23,31,62,65–67,69,71), omitting Quinley et al. (71) substantially increased the summary odds ratio (from 1.79 to 1.93). Neither the funnel plot pattern (Figure 3) nor the results from the asymmetry hypothesis tests (Begg test: $P = .17$; Egger test: $P = .54$) suggested evidence of publication bias.

Completeness of Reporting on Selected Aspects of Internal and External Validity

Nonresponse at baseline was not an issue for the nine studies (23,31,59,62,63,65,66,69,71) that used records to identify and track the study sample for outcome measurement (Table 3). For the other 16 studies, all but one (64) reported the response rate at baseline; however, of these 16 studies, only two (29,74) compared characteristics of respondents and nonrespondents at baseline. Nineteen of the 25 studies tested for equivalence of study groups at baseline on selected variables (Table 3). Fifteen of the 16 studies that actively recruited participants (as opposed to tracking them passively through medical records or administrative databases) reported response rates at follow-up; however, only four (28,58,75,76) compared the final sample with dropouts. Only eight studies that should have tested for differential attrition across study groups did so (Table 3).

Table 2. Tests of heterogeneity, odds ratios (ORs), and 95% confidence intervals (CIs) for potential explanatory variables of repeat mammography screening*

Study variable	No. of estimates	First author (reference)	Q	df	P	I ²	OR (95% CI)
All studies	27†		69.5	26	<.001	62.6	1.39 (1.27 to 1.52)
Age, y							
≥40	9	DeFrank (70), Eaker (60), Finney (23), Goel (62), Margolis (73), Partin (63), Rimer (77), Schapira (69), Simon (31)	21.9	8	.005	63.5	1.31 (1.08 to 1.59)
≥50	18	Andersen (61), Barr (65), Bodiya (66), Clark-1 (72), Clark-2 (72), Costanza-1 (75), Costanza-2 (75), Crane (58), Drossaert (59), Duan (64), Lipkus (32), Mayer (67), Messina (76), Michielutte (74), Quinley (71), Rakowski (68), Skinner (78), Vernon (29)	44.8	17	<.001	62.1	1.44 (1.29 to 1.60)
Study setting							
Health care	19	Barr (65), Bodiya (66), Clark-1 (72), Clark-2 (72), Costanza-1 (75), Costanza-2 (75), DeFrank (70), Finney (23), Lipkus (32), Margolis (73), Mayer (67), Messina (76), Michielutte (74), Quinley (71), Rakowski (68), Rimer (77), Schapira (69), Simon (31), Skinner (78)	59.9	17	<.001	71.6	1.40 (1.20 to 1.65)
Community	8	Andersen (61), Crane (58), Drossaert (59), Duan (64), Eaker (60), Goel (62), Partin (63), Vernon (29)	9.3	7	.230	25.0	1.36 (1.22 to 1.50)
Screening interval							
1 year	18	Bodiya (66), Clark-1 (72), Clark-2 (72), DeFrank (70), Duan (64), Finney (23), Goel (62), Lipkus (32), Mayer (67), Michielutte (74), Partin (63), Quinley (71), Rakowski (68), Rimer (77), Schapira (69), Simon (31), Skinner (78), Vernon (29)	48.3	17	<.001	64.8	1.41 (1.27 to 1.57)
2 years	7	Andersen (61), Barr (65), Costanza-1 (75), Costanza-2 (75), Crane (58), Drossaert (59), Messina (76)	14.6	6	.023	59.0	1.44 (1.16 to 1.78)
Age dependent	2	Eaker (60), Margolis (73)	NA	NA	NA	NA	NA
Study design							
Design 1: measured one pre- and one postintervention mammogram	20	Andersen (61), Barr (65), Bodiya (66), Clark-1 (72), Costanza-1 (75), Crane (58), DeFrank (70), Drossaert (59), Duan (64), Eaker (60), Finney (23), Goel (62), Margolis (73), Mayer (67), Michielutte (74), Partin (63), Quinley (71), Rakowski (68), Schapira (69), Simon (31)	60.9	19	<.001	68.8	1.44 (1.29 to 1.62)
Design 2: measured two postintervention mammograms	7	Clark-2 (72), Costanza-2 (75), Lipkus (32), Messina (76), Rimer (77), Skinner (78), Vernon (29)	6.64	6	.355	9.7	1.26 (1.12 to 1.42)
Mammography data source							
Medical records, administrative, or program data	14	Barr (65), Bodiya (66), Clark-1 (72), Clark-2 (72), Drossaert (59), Finney (23), Goel (62), Mayer (67), Michielutte (74), Partin (63), Quinley (71), Rakowski (68), Simon (31), Skinner (78)	43.7	13	<.001	70.3	1.42 (1.25 to 1.62)
Self-report	10	Andersen (61), Costanza-1 (75), Costanza-2 (75), Crane (58), Duan (64), Eaker (60), Lipkus (32), Messina (76), Rimer (77), Vernon (29)	9.8	9	.367	8.2	1.36 (1.23 to 1.52)
Medical record or self-report intervention strategy class 1†	3	DeFrank (70), Margolis (73), Schapira (69)	NA	NA	NA	NA	NA
Reminder	8	Barr (65), Bodiya (66), Finney (23), Goel (62), Mayer (67), Quinley (71), Schapira (69), Simon (31)	42.2	7	<.001	83.4	1.79 (1.41 to 2.29)
Education/motivation	7	Clark-1 (72), Clark-2 (72), DeFrank (70), Drossaert (59), Eaker (60), Partin (63), Vernon (29)	2.5	6	.868	0.0	1.25 (1.14 to 1.38)
Counseling	12	Andersen (61), Costanza-1 (75), Costanza-2 (75), Crane (58), Duan (64), Lipkus (32), Margolis (73), Messina (76), Michielutte (74), Rakowski (68), Rimer (77), Skinner (78)	18.6	11	.068	40.9	1.28 (1.15 to 1.43)

(Table continues)

Table 2 (continued).

Study variable	No. of estimates	First author (reference)	Q	df	P	I ²	OR (95% CI)
Intervention strategy class 2							
Barriers-specific telephone counseling							
Yes	9	Andersen (61), Costanza-1 (75), Costanza-2 (75), Crane (58), Lipkus (32), Messina (76), Michielutte (74), Rakowski (68), Rimer (77)	10.1	8	.261	20.5	1.35 (1.19 to 1.53)
No	18	Barr (65), Bodiya (66), Clark-1 (72), Clark-2 (72), DeFrank (70), Drossaert (59), Duan (64), Eaker (60), Finney (23), Goel (62), Margolis (73), Mayer (67), Partin (63), Quinley (71), Schapira (69), Simon (31), Skinner (78), Vernon (29)	59.4	17	<.001	71.4	1.40 (1.25 to 1.58)
Intervention strategy class 3							
Single	14	Barr (65), Bodiya (66), Clark-1 (72), Clark-2 (72), Goel (62), Drossaert (59), Duan (64), Finney (23), Margolis (73), Mayer (67), Quinley (71), Schapira (69), Simon (31), Vernon (29)	51.6	13	<.001	74.8	1.50 (1.30 to 1.73)
Multiple	13	Andersen (61), Costanza-1 (75), Costanza-2 (75), Crane (58), DeFrank (70), Eaker (60), Lipkus (32), Messina (76), Michielutte (74), Partin (63), Rakowski (68), Rimer (77), Skinner (78)	15.1	12	.237	20.5	1.28 (1.17 to 1.39)
Delivery mode class 1 \$							
Mail only	10	Clark-1 (72), Clark-2 (72), DeFrank (70), Drossaert (59), Finney (23), Mayer (67), Partin (63), Quinley (71), Simon (31), Vernon (29)	26.1	9	.002	65.5	1.34 (1.19 to 1.51)
Telephone only	3	Barr (65), Duan (64), Lipkus (32)	NA	NA	NA	NA	NA
Mail plus telephone	10	Bodiya (66), Costanza-1 (75), Costanza-2 (75), Crane (58), Goel (62), Messina (76), Michielutte (74), Rakowski (68), Rimer (77), Skinner (78)	20.9	9	.013	56.8	1.41 (1.16 to 1.71)
Mail plus in person	2	Margolis (73), Schapira (69)	NA	NA	NA	NA	NA
Community education plus other modes	2	Andersen (61), Eaker (60)	NA	NA	NA	NA	NA
Delivery mode class 2							
Single	13	Barr (65), Clark-1 (72), Clark-2 (72), DeFrank (70), Drossaert (59), Duan (64), Finney (23), Lipkus (32), Mayer (67), Partin (63), Quinley (71), Simon (31), Vernon (29)	35.5	12	<.001	66.2	1.41 (1.26 to 1.58)
Multiple	14	Andersen (61), Bodiya (66), Costanza-1 (75), Costanza-2 (75), Crane (58), Eaker (60), Goel (62), Margolis (73), Messina (76), Michielutte (74), Rakowski (68), Rimer (77), Schapira (69), Skinner (78)	33.8	13	<.001	61.5	1.38 (1.16 to 1.64)
Control group type							
Active	13	Barr (65), Costanza-1 (75), Costanza-2 (75), DeFrank (70), Drossaert (59), Finney (23), Goel (62), Lipkus (32), Michielutte (74), Rakowski (68), Rimer (77), Schapira (69), Simon (31)	35.0	12	<.001	65.7	1.40 (1.19 to 1.64)
Survey only	12	Andersen (61), Clark-1 (72), Clark-2 (72), Crane (58), Duan (64), Eaker (60), Margolis (73), Mayer (67), Messina (76), Partin (63), Skinner (78), Vernon (29)	29.6	11	.002	62.8	1.36 (1.16 to 1.59)
No contact	2	Bodiya (66), Quinley (71)	NA	NA	NA	NA	NA
Use of theoretic framework							
Yes	19	Andersen (61), Barr (65), Clark-1 (72), Clark-2 (72), Costanza-1 (75), Costanza-2 (75), Crane (58), DeFrank (70), Drossaert (59), Duan (64), Finney (23), Lipkus (32), Mayer (67), Messina (76), Michielutte (74), Rakowski (68), Rimer (77), Skinner (78), Vernon (29)	47.7	18	<.001	62.2	1.36 (1.21 to 1.53)
No	8	Bodiya (66), Eaker (60), Goel (62), Margolis (73), Partin (63), Quinley (71), Schapira (69), Simon (31)	21.8	7	.003	67.9	1.49 (1.22 to 1.81)

(Table continues)

Table 2 (continued).

Study variable	No. of estimates	First author (reference)	Q	df	P	I ²	OR (95% CI)
Theoretic construct							
Barriers							
Yes	14	Andersen (61), Barr (65), Costanza-1 (75), Costanza-2 (75), Crane (58), DeFrank (70), Duan (64), Lipkus (32), Messina (76), Michielutte (74), Rakowski (68), Rimer (77), Skinner (78), Vernon (29)	25.2	13	.022	48.3	1.35 (1.19 to 1.52)
No	13	Bodiya (66), Clark-1 (72), Clark-2 (72), Drossaert (59), Eaker (60), Finney (23), Goel (62), Margolis (73), Mayer (67), Partin (63), Quinley (71), Schapira (69), Simon (31)	44.1	12	<.001	72.8	1.45 (1.25 to 1.68)
Stage of change							
Yes	10	Clark-1 (72), Clark-2 (72), Crane (58), Lipkus (32), Messina (76), Michielutte (74), Rakowski (68), Rimer (77), Skinner (78), Vernon (29)	9.7	9	.376	7.2	1.31 (1.18 to 1.45)
No	17	Andersen (61), Barr (65), Bodiya (66), Costanza-1 (75), Costanza-2 (75), DeFrank (70), Drossaert (59), Duan (64), Eaker (60), Finney (23), Goel (62), Margolis (73), Mayer (67), Partin (63), Quinley (71), Schapira (69), Simon (31)	59.0	16	<.001	72.9	1.44 (1.27 to 1.64)
Use of tailoring							
Yes	14	Andersen (61), Clark-1 (72), Clark-2 (72), Costanza-1 (75), Costanza-2 (75), Crane (58), Duan (64), Lipkus (32), Messina (76), Michielutte (74), Rakowski (68), Rimer (77), Skinner (78), Vernon (29)	14.7	13	.325	11.7	1.32 (1.21 to 1.44)
No	13	Barr (65), Bodiya (66), DeFrank (70), Drossaert (59), Eaker (60), Finney (23), Goel (62), Margolis (73), Mayer (67), Partin (63), Quinley (71), Schapira (69), Simon (31)	54.1	12	<.001	77.8	1.46 (1.26 to 1.70)

* Odds ratios and 95% confidence intervals were computed under random-effects model assumptions for heterogeneous subgroups and under fixed-effects model assumptions for homogeneous subgroups. A statistically significant Q indicates a heterogeneous distribution of study effect sizes, which may then warrant additional subgroup analyses (30). Statistical tests were two-sided. Categories of a variable were rated NA if there were fewer than five estimates; heterogeneity tests were not performed for these categories. I² = percentage of total variation across studies that is because of heterogeneity rather than chance (33); NA = not applicable; Q = the Cochran Q statistic.

† Two of the 25 studies, Clark et al. (72) and Costanza et al. (75), provided two estimates, so the total number of estimates was 27. We designated these estimates 1 and 2.

‡ Some studies used more than one intervention strategy, for example, reminder plus counseling. We classified those studies based on the most intensive strategy used. Intervention strategies were classified and analyzed in three different ways.

§ Delivery modes were classified and analyzed in two different ways.

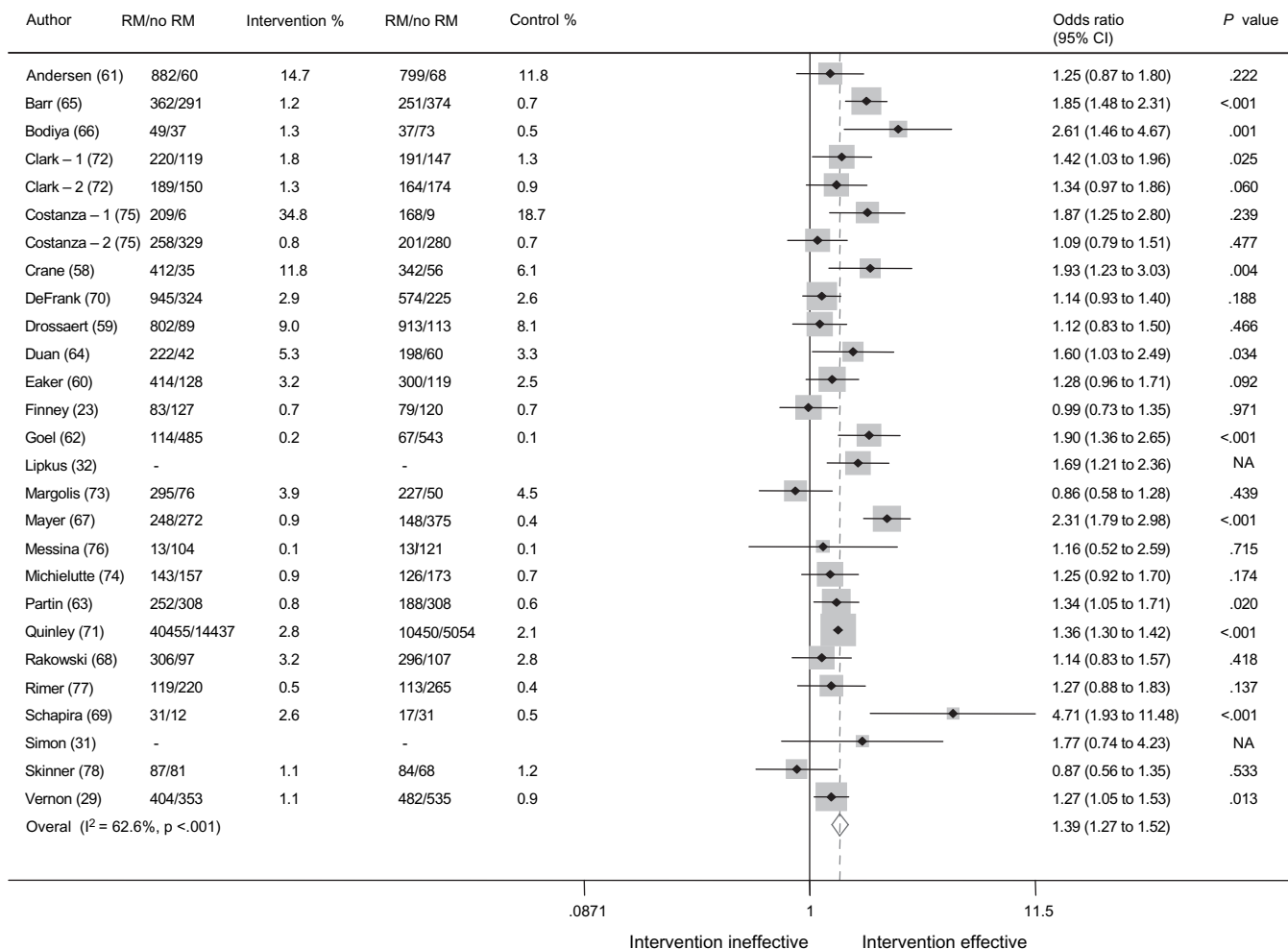


Figure 2. Forest plot comparing intervention groups to control groups on repeat mammography by author. Sizes of study-specific odds ratios (ORs) (in gray) are inversely proportional to the variance. The 95% confidence intervals (CIs) for the overall summary odds ratio were computed under random-effects model assumptions. Weights are from the

random-effects analysis. Statistical tests were two-sided. Control % = odds of RM compared with odds of no RM in the control group; intervention % = odds of RM compared with odds of no RM in the intervention group; NA = not available; OR = ratio of the intervention odds to the control odds; RM/no RM = repeat mammography/no repeat mammography.

Only six studies reported how the sample size for the intervention trial was determined, and only 11 studies conducted intent-to-treat analyses (Table 3), defined as including everyone who was randomized to an intervention or control group (79). Of the four group-level, randomized trials (61,64,74,75), all described analyses intended to adjust for the design effect; however, neither of the two group-level, nonrandomized trials (60,71) adjusted for the effect of nested samples. Two of the studies (67,76) that identified individuals as the unit of randomization and analysis could have but did not report the intraclass correlation or magnitude of the design effect to describe the possible influence of patients nested within providers or facilities.

Discussion

We observed statistically significant heterogeneity ($P < .001$) among the effect size estimates (ORs) in the 25 studies (27 estimates), indicating variability in effect estimates of repeat mammography rates. Of the 15 categorical covariates identified a priori as study characteristics

that may influence effect size, no single covariate resolved the heterogeneity in univariate analyses. In multivariable meta-regression, only the intervention strategy of reminders vs the combined categories of education/motivation and counseling remained a statistically significant predictor of the magnitude of the intervention effect as measured by the odds ratio; however, substantial residual heterogeneity persisted in the model. The summary odds ratio for the eight heterogeneous studies using reminders was the largest observed (OR = 1.79, 95% CI = 1.41 to 2.29 computed under a random-effects model) and was statistically significantly ($P_{diff} = .008$) greater than the summary odds ratio for the homogeneous group of 17 studies that used the more intensive strategies of education/motivation or counseling (OR = 1.27, 95% CI = 1.17 to 1.37) regardless of whether it was computed under a fixed- or random-effects model. It is important to note that the eight studies using reminders showed statistically significant heterogeneity despite the notoriously low statistical power of homogeneity testing (80). Moreover, all eight studies were alike in terms of using medical records or administrative data to ascertain mammography status using design 1 (one pre- and one postintervention

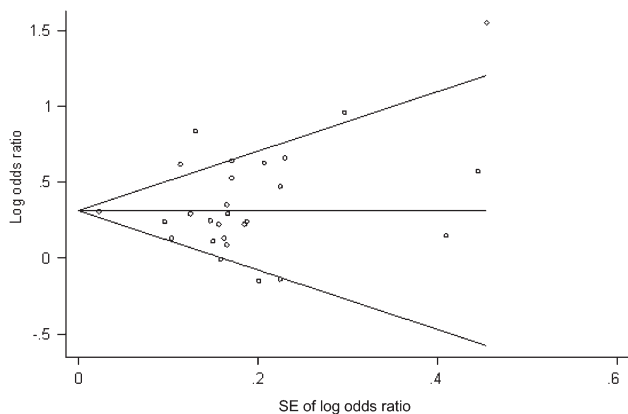


Figure 3. Begg funnel plot of all studies. The Begg test (39) is a measure of publication bias and tests whether the Begg rank correlation between effect size and its SE is zero. Pseudo confidence intervals are the **points** connected by the **diagonal lines** forming the “funnel” on the funnel plot; they are the expected 95% confidence intervals for a given SE (depicted as increasing along the x-axis).

mammogram) and, except for one study, being conducted within a health-care setting. The results of the influence analyses confirmed that the observed heterogeneity was mostly attributable to one or two of the studies using reminders (65,67). Because of this heterogeneity, we cannot conclude that the use of a reminder intervention strategy within a health-care setting is more effective than alternate intervention strategies in the same or different study settings. Therefore, additional studies are needed to help resolve the remaining heterogeneity in this subgroup by identifying the explanatory study characteristics or research methodologies that are the key factors in increasing repeat mammography.

The 17 studies that used education/motivation or counseling were remarkably homogeneous in their effect sizes with a narrow confidence interval, suggesting a high degree of consistency among the studies and that the true intervention effect of these strategies may be, at best, moderate, that is, odds ratios between 1.18 and 1.36. The results of the meta-regression modeling further suggest that, among these homogeneous studies, there was no detectable advantage or disadvantage in the different study designs, methods, settings, populations, intervention strategies, delivery modes, outcome measurements, screening intervals, or use of theory. This finding raises a question as to whether substantial increases in regular mammography screening can be expected from education/motivation or counseling interventions, regardless of how intensive, rigorous, innovative, or expensive the approach. In other words, changes in regular mammography screening behavior may not be particularly sensitive to variations in education/motivation or counseling interventions. In the current US environment, substantial increases in regular cancer screening behavior may depend more on factors at the systems level (eg, regulations relating to health-care access such as insurance coverage and standards of preventive care) than on factors at the individual level such as perceived risk of breast cancer.

Our finding of a relatively modest intervention effect for the subgroup of more intensive intervention studies is consistent with the finding in a meta-analysis (16) that was restricted to tailored interventions and that focused on one-time mammography

screening. Sohl and Moyer (16) found that interventions promoting repeat mammography had a smaller effect size (OR = 1.17) compared with those promoting one-time use (OR = 1.53). As discussed previously, however, we cannot conclude that lower intensity interventions such as reminders are a better strategy without additional research to determine whether a particular type of reminder strategy is effective across different study settings, populations, and study methodologies.

An unexpected finding was that there were two types of study designs used in repeat mammography interventions. In design 1, women who had had a recent preintervention mammogram at study baseline were followed long enough to complete one on-schedule mammogram during the study period. In design 2, women with diverse mammography histories at study baseline were followed long enough to complete two on-schedule mammograms during the study period. In design 2 studies, women who were overdue or had never been screened may have been more resistant to attempts to get them to complete screening. Although the confidence intervals overlapped, the odds ratio for design 2 studies was smaller compared with design 1 studies, suggesting that women who were not overdue at study baseline were more likely to complete another mammogram on schedule compared with a group of women that included some who were overdue. Future studies should consider how mammography history may affect receptivity to different types of interventions (81). For example, if a woman experiences the procedure as painful, she may be unwilling to return for her next mammogram when it is due and may disregard reminders or messages promoting mammography.

There is no consensus about how to classify types of intervention strategies. Systematic reviews and meta-analyses of one-time mammography screening have used a number of different classifications (7–19). The lack of consistency may result, in part, from lack of consensus about how to operationalize our theoretic frameworks and constructs (24). In addition, many of the interventions reviewed here were complex, multicomponent interventions and, therefore, difficult to classify. For these reasons, we explored three approaches to classifying intervention strategies. None of these approaches yielded homogeneous subgroups across all categories of a variable, and, with the possible exception of reminders, the effect sizes were generally similar.

In our assessment of the quality of reporting for eight characteristics of internal validity, most studies tested for equivalence of study groups at baseline and reported the response rate at follow-up (Table 3). Fewer studies reported whether they compared the characteristics of participants who remained in the study with those who dropped out, and even fewer reported whether there was differential attrition by study group (Table 3). Only 11 studies conducted an intention-to-treat analysis, so it is likely that the effect sizes were overestimated because data on other cancer screening behaviors suggest that dropouts are less likely to complete screening tests (82–84). There was far less attention to reporting study characteristics that affect external validity such as the representativeness of participants and settings. As noted by Steckler and McElroy (85), “Systematic reviews and meta-analyses are limited in the conclusions that can be drawn when external validity data are not reported.” This limitation needs to be addressed if we are to successfully disseminate effective interventions. Application

Table 3. Completeness of reporting on selected aspects of internal and external validity*

Study	Preintervention			Postintervention/follow-up			Analysis			
	Response rate at baseline	Baseline comparison of respondents and nonrespondents	Test for equivalence of study groups at baseline	Response rate at follow-up	Compare final sample and dropouts	Test for differential attrition by study group	Sample size determination (a priori statistical power analysis)	Intent-to-treat	Unit of randomization or assignment	For nested samples: appropriate analysis/adjustment
Andersen (61)	Y	N	N	Y	N	N	N	N	Group (community)	Y
Barr (65)	NA	NA	Y	NA	NA	NA	N	Y	Individual	NA
Bodiya (66)	NA	NA	N	NA	NA	NA	N	Y	Individual	NA
Clark (72)	Y	N	N	N	N	Y	N	N	Individual	NA
Costanza (75)	Y	N	N	Y	Y	Y	N	N	Group (practices)	Y
Crane (58)	Y	N	Y	Y	Y	Y	N	N	Residences	NA
DeFrank (70)	Y	N	Y	Y	N	N	Y	N	Individual	NA
Drossaert (59)	NA	NA	Y	NA	NA	NA	N	Y	Individual	NA
Duan (64)	N	N	Y	Y	N	N	N	N	Group (church)	Y
Eaker (60)	Y	N	Y	Y	N	N	N	N	Group (county)†	N
Finney (23)	NA	NA	N	NA	NA	NA	Y	Y	Individual	NA
Goel (62)	NA	NA	Y	NA	NA	NA	N	Y	Individual	NA
Lipkus (32)	Y	N	Y	Y	N	N	N	N	Individual	NA
Margolis (73)	Y	N	Y	Y	N	Y	N	Y	Individual	NA
Mayer (67)	Y	N	Y	NA	NA	N	Y	Y	Individual	NA
Messina (76)	Y	N	Y	Y	Y	N	N	N	Individual	NA
Michielutte (74)	Y	Y	Y	Y	N	Y	N	N	Group (practices)	Y
Partin (63)	NA	NA	Y	NA	NA	NA	N	Y	Individual	NA
Quinley (71)	NA	NA	Y	NA	NA	NA	N	Y	Group (mammography facilities)‡	N
Rakowski (68)	Y	N	Y	Y	N	N	Y	N	Individual	NA
Rimer (77)	Y	N	Y	Y	N	N	N	N	Individual	NA
Schapira (69)	NA	NA	Y	Y	N	Y	N	N	Individual	NA
Simon (31)	NA	NA	Y	NA	NA	NA	Y	Y	Individual	NA
Skinner (78)	Y	N	N	Y	N	Y	N	N	Individual	NA
Vernon (29) and del Junco (28)	Y	Y	Y	Y	Y	Y	Y	Y	Individual	NA

* Studies coded NA used a data source such as medical records to identify the study population, and the investigators were able to enumerate and track all eligible women; therefore, issues related to response rates, equivalence, and follow-up were not a problem. N = no; NA = not applicable; Y = yes.

† We defined intent-to-treat analysis as including all eligible randomized persons in the outcome analysis whether or not they completed the intervention and/or all study surveys (79). For example, if participants were randomized to a study group but were later excluded from the outcome analysis because they did not complete follow-up surveys, a study was rated “N.”

‡ Eaker et al. (60) and Quinley et al. (71) used quasi-experimental designs. Eaker et al. (60) assigned counties (not randomized) to intervention or control group. Quinley et al. (71) compared mammography facilities that mailed reminders with those that did not.

of frameworks to address internal (41,86) and external (44,87) validity in the implementation and evaluation of interventions will enable us to learn from our successes as well as our failures.

A limitation of our systematic review is that several of the studies reviewed here were not explicitly designed to promote repeat mammography, although they provided data that allowed us to calculate effect estimates. It may be that had their interventions been designed to address repeat mammography in addition to one-time mammography, the effect estimates in those studies would have been different. Although there is an extensive body of intervention research on one-time mammography screening, the number of intervention studies of repeat mammography is comparatively small, and estimates in some categories of the predictor variables were unstable. In addition, most of the studies were conducted more than 10 years ago, and most were of non-Hispanic white women, thus limiting our ability to generalize the findings to the present day and to other ethnic groups.

If we are to reap the benefits of mortality reduction from mammography screening, we need a better understanding of the determinants of repeat screening behavior so that we can develop more effective interventions. This review called attention to a number of characteristics in the studies to date, which, if attended to in future studies, will increase our understanding of how to develop and implement interventions to increase regular mammography screening. Because reporting standards are increasingly being adopted by journal editors, it will be easier to synthesize the literature and draw conclusions about what works, under what circumstances, and for what reasons.

References

- Jemal A, Siegel R, Ward E, et al. Cancer statistics. *CA Cancer J Clin*. 2008;58(2):71–96.
- Kerlikowske K, Grady D, Rubin SM, Sandrock C, Ernster VL. Efficacy of screening mammography: a meta-analysis. *JAMA*. 1995;273(2):149–154.
- U.S. Preventive Services Task Force (USPSTF). *Screening for Breast Cancer. Recommendations and Rationale*. Rockville, MD: U.S. Department Health and Human Services; 2002. <http://www.ahrq.gov/clinic/uspstf09/breastcancer/brcanrs.htm>. Accessed May 11, 2009.
- Swan J, Breen NL, Coates RJ, Rimer BK, Lee NC. Progress in cancer screening practices in the United States: results from the 2000 National Health Interview Survey. *Cancer*. 2003;97(6):1528–1540.
- Breen N, Cronin KA, Meissner HI, et al. Reported drop in mammography: is this cause for concern. *Cancer*. 2007;109(12):2405–2409.
- Clark MA, Rakowski W, Bonacore LB. Repeat mammography: prevalence estimates and considerations for assessment. *Ann Behav Med*. 2003;26(3):201–211.
- Bonfill CX, Marzo CM, Pladevall VM, Marti J, Empananza JI. Strategies for increasing the participation of women in community breast cancer screening. *Cochrane Database Syst Rev* 2001; (1): Art. No.:CD002943. doi: 10.1002/14651858.CD002943.
- Legler JM, Meissner HI, Coyne C, Breen NL, Chollette V, Rimer BK. The effectiveness of interventions to promote mammography among women with historically lower rates of screening. *Cancer Epidemiol Biomarkers Prev*. 2002;11(1):59–71.
- Ratner PA, Bottorff JL, Johnson JL, Cook R, Lovato CY. A meta-analysis of mammography screening promotion. *Cancer Detect Prev*. 2001;25(3):147–160.
- Wagner TH. The effectiveness of mailed patient reminders on mammography screening: a meta-analysis. *Am J Prev Med*. 1998;14(1):64–70.
- Yabroff KR, Mandelblatt JS. Interventions targeted toward patients to increase mammography use. *Cancer Epidemiol Biomarkers Prev*. 1999;8(9):749–757.
- Yabroff KR, O'Malley AS, Mangan P, Mandelblatt JS. Inreach and outreach interventions to improve mammography use. *J Am Med Womens Assoc*. 2001;56(4):166–174.
- Mandelblatt JS, Yabroff KR. Effectiveness of interventions designed to increase mammography use: a meta-analysis of provider-targeted strategies. *Cancer Epidemiol Biomarkers Prev*. 1999;8(9):759–767.
- Stone EG, Morton SC, Hulscher MEJL, et al. Interventions that increase use of adult immunization and cancer screening services: a meta-analysis. *Ann Intern Med*. 2002;136(9):641–651.
- Snell JL, Buck EL. Increasing cancer screening: a meta-analysis. *Prev Med*. 1996;25(6):702–707.
- Sohl SJ, Moyer A. Tailored interventions to promote mammography screening: a meta-analysis review. *Prev Med*. 2007;45(4):252–261.
- Baron RC, Rimer BK, Coates RC, et al. Client-directed interventions to increase community access to breast, cervical, and colorectal cancer screening. *Am J Prev Med*. 2008;35(suppl 1):S56–S66.
- Sabatino SA, Habarta N, Baron RC, et al. Interventions to increase recommendation and delivery of screening for breast, cervical, and colorectal cancers by healthcare providers: systematic reviews of provider assessment and feedback and provider incentives. *Am J Prev Med*. 2008;35(suppl 1):S67–S74.
- Baron RC, Rimer BK, Breslow RA, et al. Client-directed interventions to increase community demand for breast, cervical, and colorectal cancer screening. *Am J Prev Med*. 2008;35(1 suppl):S34–S55.
- Yabroff KR. Interventions to improve cancer screening: commentary from a health services research perspective. *Am J Prev Med*. 2008; 35(suppl 1):S6–S9.
- Higgins JPT, Green S, eds. *Cochrane Handbook for Systematic Reviews of Interventions Version 5.0.2 [updated September 2009]. The Cochrane Collaboration*. 2009. <http://www.cochrane-handbook.org>. Accessed June 1, 2010.
- Janz NK, Becker MH. The health belief model: a decade later. *Health Educ Q*. 1984;11(1):1–47.
- Finney LJ, Iannotti RJ. Message framing and mammography screening: a theory-driven intervention. *Behav Med*. 2002;28(1):5–14.
- Vernon SW, Tiro JA, Meissner HI. Behavioral research in cancer screening. In: Miller SM, Bowen DJ, Croyle RT, Roland JH, eds. *Handbook of Cancer Control and Behavioral Science. A Resource for Researchers, Practitioners, and Policymakers*. 1st ed. Washington, DC: American Psychological Association; 2009:255–278.
- Kreuter MW, Strecher VJ, Glassman B. One size does not fit all: the case for tailoring print materials. *Ann Behav Med*. 1999;21(4):276–283.
- Kreuter MW, Wray RJ. Tailored and targeted health communications: strategies for enhancing information relevance. *Am J Health Behav*. 2003;27(suppl 3):S227–S232.
- The NCI Breast Cancer Screening Consortium. Screening mammography: a missed clinical opportunity? The results of the NCI Breast Cancer Screening Consortium and National Health Interview Survey Studies. *JAMA*. 1990;264(1):54–58.
- del Junco DJ, Vernon SW, Coan SP, et al. Promoting regular mammography screening: I. A systematic assessment of validity in a randomized trial. *J Natl Cancer Inst*. 2008;100(5):333–346.
- Vernon SW, del Junco DJ, Tiro JA, et al. Promoting regular mammography screening: II. Results from a randomized controlled trial in U.S. women veterans. *J Natl Cancer Inst*. 2008;100(5):347–358.
- Lipsey MW, Wilson DB. The way in which intervention studies have “personality” and why it is important to meta-analysis. *Eval Health Prof*. 2001;24(3):236–254.
- Simon MS, Gimotty PA, Moncrease A, Dews P, Burack RC. The effect of patient reminders on the use of screening mammography in an urban health department primary care setting. *Breast Cancer Res Treat*. 2001; 65(1):63–70.
- Lipkus IM, Rimer BK, Halabi S, Strigo TS. Can tailored interventions increase mammography use among HMO women? *Am J Prev Med*. 2000; 18(1):1–10.
- Higgins JP, Thompson SG, Deeks JJ, Altman DG. Measuring inconsistency in meta-analyses. *BMJ*. 2003;327(7414):557–560.
- StataCorp. *Stata Statistical Software: Release 10*. College Station, TX: StataCorp LP; 2007.

35. Thompson SG, Higgins JP. How should meta-regression analyses be undertaken and interpreted? *Stat Med.* 2002;21(11):1559–1573.
36. Sharp SI. sbe23: Meta-analysis regression. *Stata Tech Bull.* 1998;7(42):16–22.
37. Diclemente CC, Prochaska JO. Toward a comprehensive transtheoretical model of change. In: Miller WR, Heather N, eds. *Treating Addictive Behaviors.* 2nd ed. New York, NY: Plenum Press; 2002:3–24.
38. Tobias A. sbe26: Assessing the influence of a single study in meta-analysis. *Stata Tech Bull.* 1999;8(47):15–17.
39. Begg CB, Mazumdar M. Operating characteristics of a rank correlation test for publication bias. *Biometrics.* 1994;50(4):1088–1101.
40. Egger M, Smith GD, Schneider M, Minder C. Bias in meta-analysis detected by a simple, graphical test. *BMJ.* 1997;315(7):629–634.
41. Kaplan RM, Trudeau KJ, Davidson KW. New policy on reports of randomized clinical trials [editorial]. *Ann Behav Med.* 2004;27(2):81.
42. Murray DM, Pals SL, Blitstein JL, Alfano CM, Lehman J. Design and analysis of group-randomized trials in cancer: a review of current practices. *J Natl Cancer Inst.* 2008;100(7):483–491.
43. Glasgow RE, Vogt TM, Boles SM. Evaluating the public health impact of health promotion interventions: the RE-AIM framework. *Am J Public Health.* 1999;89(9):1322–1327.
44. Green LW, Glasgow RE. Evaluating the relevance, generalization, and applicability of research: issues in external validation and translation methodology. *Eval Health Prof.* 2006;29(1):126–153.
45. Bird JA, McPhee SJ, Ha N-T, Le B, Davis T, Jenkins CNH. Opening pathways to cancer screening for Vietnamese-American women: lay health workers hold a key. *Prev Med.* 1998;27(6):821–829.
46. Burack RC, Gimotty PA. Promoting screening mammography in inner-city settings: the sustained effectiveness of computerized reminders in a randomized controlled trial. *Med Care.* 1997;35(9):921–931.
47. Foley EC, D'Amico F, Merenstein JH. Five-year follow-up of a nurse-initiated intervention to improve mammography recommendation. *J Am Board Fam Pract.* 1995;8(6):452–456.
48. Otero-Sabogal R, Owens D, Canchola J, Tabnak F. Improving rescreening in community clinics: does a system approach work? *J Community Health.* 2006;31(6):497–519.
49. Ruffin MT, Gorenflo DW. Interventions fail to increase cancer screening rates in community-based primary care practices. *Prev Med.* 2004;39(3):435–440.
50. Abood DA, Coster DC, Mullis AK, Black DR. Evaluation of a “loss-framed” minimal intervention to increase mammography utilization among medically un- and under-insured women. *Cancer Detect Prev.* 2002;26(5):394–400.
51. Carney PA, Harwood BG, Greene MA, Goodrich ME. Impact of a telephone counseling intervention on transitions in stage of change and adherence to interval mammography screening (United States). *Cancer Causes Control.* 2005;16(7):799–807.
52. Preston JA, Scinto JD, Grady JN, Schulz AF, Petrillo MK. The effect of a multifaceted physician office-based intervention on older women's mammography use. *J Am Geriatr Soc.* 2000;48(1):1–7.
53. Partin MR, Slater JS, Caplan L. Randomized controlled trial of a repeat mammography intervention: effect of adherence definitions on results. *Prev Med.* 2005;41(3):734–740.
54. Prochaska JO, Velicer WF, Redding CA, et al. Stage-based expert systems to guide a population of primary care patients to quit smoking, eat healthier, prevent skin cancer, and receive regular mammograms. *Prev Med.* 2005;41(2):406–416.
55. Stoddard AM, Fox SA, Costanza ME, et al. Effectiveness of telephone counseling for mammography: results from five randomized trials. *Prev Med.* 2002;34(1):90–99.
56. Bloom JR, Stewart SL, Hancock SL. Breast cancer screening in women surviving Hodgkin disease. *Am J Clin Oncol.* 2006;29(3):258–266.
57. Richardson JL, Mondrus GT, Danley K, Deapen D, Mack T. Impact of a mailed intervention on annual mammography and physician breast examinations among women at high risk of breast cancer. *Cancer Epidemiol Biomarkers Prev.* 1996;5(1):71–76.
58. Crane LA, Leakey TA, Rimer BK, Wolfe P, Woodworth MA, Warnecke RB. Effectiveness of a telephone outcall intervention to promote screening mammography among low-income women. *Prev Med.* 1998;27(5, pt 2):1–49.
59. Drossaert CHC, Boer H, Seydel ER. Health education to improve repeat participation in the Dutch Breast Cancer Screening Programme—evaluation of a leaflet tailored to previous participants. *Patient Educ Couns.* 1996;28(2):121–131.
60. Eaker ED, Jaros L, Vierkant RA. Women's Health Alliance Intervention Study: increasing community breast and cervical cancer screening. *J Public Health Manag Pract.* 2001;7(5):20–30.
61. Andersen MR, Yasui Y, Meischke H, Kuniyuki A, Etzioni RD, Urban N. The effectiveness of mammography promotion by volunteers in rural communities. *Am J Prev Med.* 2000;18(3):199–207.
62. Goel A, George J, Burack RC. Telephone reminders increase re-screening in a county breast screening program. *J Health Care Poor Underserved.* 2008;19(2):512–521.
63. Partin MR, Malone M, Winnett M, Slater J, Bar-Cohen A, Caplan LS. The impact of survey non-response bias on conclusions drawn from a mammography intervention trial. *J Clin Epidemiol.* 2003;56(9):867–873.
64. Duan N, Fox SA, Derosé KP, Carson S. Maintaining mammography adherence through telephone counseling in a church-based trial. *Am J Public Health.* 2000;90(9):1468–1471.
65. Barr JK, Franks AL, Lee NC, Antonucci DM, Rifkind S, Schachter M. A randomized intervention to improve ongoing participation in mammography. *Am J Manag Care.* 2001;7(9):887–894.
66. Bodiya A, Vorias D, Dickson HA. Does telephone contact with a physician's office staff improve mammogram screening rates? *Fam Med.* 1999;31(5):324–326.
67. Mayer JA, Lewis EC, Slymen DJ, et al. Patient reminder letters to promote annual mammograms: a randomized controlled trial. *Prev Med.* 2000;31(4):315–322.
68. Rakowski W, Lipkus IM, Clark MA, et al. Reminder letter, tailored stepped-care, and self-choice comparison for repeat mammography. *Am J Prev Med.* 2003;25(4):308–314.
69. Schapira DV, Kumar NB, Clark RA, Yag C. Mammography screening credit card and compliance. *Cancer.* 1992;70(2):509–512.
70. DeFrank JT, Rimer BK, Gierisch JM, Bowling JM, Farrell D, Skinner CS. Impact of mailed and automated telephone reminders on receipt of repeat mammograms: a randomized controlled trial. *Am J Prev Med.* 2009;36(6):459–467.
71. Quinley J, Mahotiere T, Messina CR, Lee T-K, Mikail C. Mammography-facility-based patient reminders and repeat mammograms for Medicare in New York State. *Prev Med.* 2004;38(1):20–27.
72. Clark MA, Rakowski W, Ehrich B, et al. The effect of a stage-matched and tailored intervention on repeat mammography. *Am J Prev Med.* 2002;22(1):1–7.
73. Margolis KL, Lurie N, McGovern PG, Tyrrell M, Slater JS. Increasing breast and cervical cancer screening in low-income women. *J Gen Intern Med.* 1998;13(8):515–521.
74. Michielutte R, Sharp PC, Foley KL, et al. Intervention to increase screening mammography among women 65 and older. *Health Educ Res.* 2005;20(2):149–162.
75. Costanza ME, Stoddard AM, Luckmann R, White MJ, Spitz-Avrinin J, Clemow L. Promoting mammography: results of a randomized trial of telephone counseling and a medical practice intervention. *Am J Prev Med.* 2000;19(1):39–46.
76. Messina CR, Lane DS, Grimson R. Effectiveness of women's telephone counseling and physician education to improve mammography screening among women who underuse mammography. *Ann Behav Med.* 2002;24(4):279–289.
77. Rimer BK, Halabi S, Skinner CS, et al. Effects of a mammography decision-making intervention at 12 and 24 months. *Am J Prev Med.* 2002;22(4):247–257.
78. Skinner CS, Kobrin SC, Monahan PO, et al. Tailored interventions for screening mammography among a sample of initially non-adherent women: when is a booster dose important? *Patient Educ Couns.* 2007;65(1):87–94.
79. Hollis S, Campbell F. What is meant by intention-to-treat analysis? Survey of published randomised controlled trials. *BMJ.* 1999;319(670):674.
80. Petitti DB. *Meta-Analysis, Decision Analysis, and Cost-Effectiveness Analysis: Methods for Quantitative Synthesis in Medicine.* 2nd ed. New York, NY: Oxford University Press, Inc; 2000.

81. Rakowski W, Breslau ES. Perspectives on behavioral and social science research on cancer screening. *Cancer*. 2004;101(suppl 5):1118–1130.
82. Hoogewerf PE, Hislop TG, Morrison BJ, Burns SD, Sizto R. Health belief and compliance with screening for fecal occult blood. *Soc Sci Med*. 1990;30(6):721–726.
83. Lindholm E, Berglund B, Haglund E, Kewenter J. Factors associated with participation in screening for colorectal cancer with faecal occult blood testing. *Scand J Gastroenterol*. 1995;30(2):171–176.
84. Kelly RB, Shank JC. Adherence to screening flexible sigmoidoscopy in asymptomatic patients. *Med Care*. 1992;30(11):1029–1042.
85. Steckler A, McElroy KR. The importance of external validity. *Am J Public Health*. 2008;98(1):9–10.
86. Moher D, Schultz KF, Altman DG. The CONSORT statement: revised recommendations for improving the quality of reports of parallel-group randomized trials. *Ann Intern Med*. 2001;134(8):657–662.
87. Glasgow RE, Green LW, Klesges LM, et al. External validity: we need to do more. *Ann Behav Med*. 2006;31(2):105–108.

Funding

National Cancer Institute (NCI)/National Institutes of Health (RO1-CA-76330 to S.W.V., D.J.d.J., and J.A.T.); NCI (RO3-CA-103512 to D.J.d.J. and S.W.V.);

UL1 (RR024148—CTSA to D.J.d.J.); NCI (R25-CA-57712 to A.M.); American Cancer Society Mentored Research Scholar Grant (CPPB-113777 to A.M.).

Notes

The funders did not have any involvement in the design of the study; the collection, analysis, and interpretation of the data; the writing of the manuscript; or the decision to submit the manuscript for publication.

The authors thank Sharon Coan for her assistance with data analysis.

Affiliations of authors: Center for Health Promotion and Prevention Research, Division of Health Promotion and Behavioral Sciences, University of Texas School of Public Health, Houston, TX (SWV); Division of Health Behavior Research, Department of Medicine, Washington University School of Medicine, St Louis, MO (AM); Division of Behavioral and Communication Sciences, Department of Clinical Sciences, University of Texas Southwestern Medical Center, Dallas, TX (JAT); Department of Surgery and Pediatrics, University of Texas Medical School and Division of Epidemiology, Human Genetics and Environmental Sciences, University of Texas School of Public Health, Houston, TX (D.J.d.J.).