



Published in final edited form as:

J Natl Cancer Inst. 2008 March 5; 100(5): 333–346. doi:10.1093/jnci/djn027.

Promoting Regular Mammography Screening I. A Systematic Assessment of Validity in a Randomized Trial

Deborah J. del Junco, Sally W. Vernon, Sharon P. Coan, Jasmin A. Tiro, Lori A. Bastian, Lara S. Savas, Catherine A. Perz, David R. Lairson, Wen Chan, Cynthia Warrick, Amy McQueen, and William Rakowski

Center for Clinical and Translational Sciences, University of Texas Health Science Center at Houston, Houston, TX (DJdJ); Division of Health Promotion and Behavioral Sciences, University of Texas-Houston School of Public Health, Houston, TX SWV, SPC, DRL, WC, AM); Cancer Prevention Fellowship Program, Office of Preventive Oncology, Division of Cancer Control and Population Sciences, National Cancer Institute, Bethesda, MD (JAT); Center for Health Services Research in Primary Care, Durham VA/Duke University, Durham, NC (LAB); Section of General Medicine, Department of Medicine, Baylor College of Medicine, Houston, TX (LSS); Department of Psychology, University of Houston-Victoria, Victoria, TX (CAP); School of Math, Science and Technology, Elizabeth City State University, Elizabeth City, NC (CW); Department of Community Health, Brown University, Providence, RI (WR).

Abstract

Background—Most health promotion trials in cancer screening offer limited evidence of external validity. We assessed internal and external validity in a nationwide, population-based trial of an intervention to promote regular mammography screening.

Methods—Beginning in September 2000, study candidates age 52 years and older ($n = 23\,000$) were randomly selected from the National Registry of Women Veterans and sent an eligibility survey. Consistent with intention-to-treat principles for effectiveness trials, we randomly assigned eligible respondents and nonrespondents to one of five groups. We mailed baseline surveys to groups 1–3 followed by intervention materials of varying personalization to groups 1 and 2. We delayed mailing baseline surveys to two additional control groups to coincide with the mailing of postintervention follow-up surveys to groups 1–3 at year 1 (group 4) and year 2 (group 5). Mammography rates were determined from self-report and Veterans Health Administration records. To assess internal validity, we compared groups on participation and factors associated with mammography screening at each stage. To assess external validity, we compared groups 3, 4, and 5 on mammography rates at the most recent follow-up to detect any cueing effects of prior surveys and at the respective baselines to uncover any secular trends. We also compared nonparticipants with participants on factors associated with mammography screening at the trial's end.

Results—We established study eligibility for 21 340 (92.8%) of the study candidates. Groups 1–3 were similar throughout the trial in participation and correlates of mammography screening. No statistically significant survey cueing effects or differences between nonparticipants and participants across groups were observed. Mammography screening rates over the 30 months preceding the respective baselines were lower in group 5 (82.3% by self-report) than in groups 1–4 (85.1%, $P = .024$, group 5 vs groups 1–4 combined), suggesting a decline over time similar to that reported for US women in general.

© The Author 2008. Published by Oxford University Press. All rights reserved.

Correspondence to: Deborah J. del Junco, PhD, Center for Clinical and Translational Sciences, University of Texas Health Science Center at Houston, 6410 Fannin St, LL.125, Houston, TX 77030 (deborah.j.deljunco@uth.tmc.edu).

Conclusion—This systematic assessment provides evidence of the trial’s internal and external validity and illustrates an approach to evaluating validity that is readily adaptable to future trials of behavioral interventions.

Most health promotion trials of interventions to change behavior evaluate efficacy rather than effectiveness (1). Efficacy trials are typically limited in scope, focusing on internal validity under optimal conditions (2). Internal validity refers to the comparability of the intervention and control groups and the extent to which their distributions on important confounding variables are equal or balanced to support a causal interpretation of any postintervention between-group differences on the study outcome of interest (2). In contrast to efficacy trials, effectiveness trials reach more broadly to a specific population, seeking external validity under real-world circumstances (3). External validity refers to the extent to which study findings can be generalized or extrapolated to the target population of interest (2). As emphasized in a recent report by the Agency for Healthcare Research and Quality (4), efficacy and effectiveness trials exist along a continuum defined by the target population and study setting, the strictness of eligibility and exclusion criteria, the health outcome measures of primary interest, the intervention regimen and mode of delivery, the adequacy of the sample size and duration of follow-up to detect minimally important differences, and the use of intention-to-treat analyses to increase the generalizability of intervention effect estimates.

The development of behavior change interventions that can be broadly disseminated is a high priority of the National Institutes of Health (5). For interventions to achieve their maximum public health impact, results from intervention trials must be generalizable (6–9). A 2004 supplement of the journal *Cancer*, jointly sponsored by the National Cancer Institute, American Cancer Society, and Centers for Disease Control and Prevention, urged researchers involved in cancer screening interventions to pay more attention to external validity (such as participant representativeness) in the design of efficacy trials (10). The supplement further emphasized that “The sooner researchers and funding agencies understand the constraints on intervention delivery in the real world, the more likely programs are to be adopted” (10).

Efficacy trials typically exclude from their analyses nonrespondents, those who refuse to participate, those who withdraw from the study, and participants with any missing data on the outcome variable of interest (11). In addition, intervention trials are seldom population based and are often restricted to a specific geographic location, health-care system or provider, and demographic profile of study participants (4). Thus, the internal and external validity of study findings and the potential for broad dissemination of most health promotion interventions remain uncertain. This uncertainty further challenges the validity of the cost-effectiveness analyses that form the basis of public and private health-care policy (12).

We sought to overcome limitations in internal and external validity by targeting the population of US women veterans in a patient-directed, nationwide intervention trial to promote regular mammography screening, Project HOME (Healthy Outlook on the Mammography Experience) (13). Guided by the CONSORT statement (14), the Code of Professional Ethics and Practices of the American Association for Public Opinion Research (15,16), the Evidence-Based Behavioral Medicine Committee (17), the Agency for Healthcare Research and Quality’s criteria for distinguishing effectiveness from efficacy trials (4), and criteria specifically recommended to improve external validity in health promotion research by Glasgow et al. (11), we systematically assessed Project HOME’s internal and external validity. Our approach included the following elements: 1) flow diagrams showing the passage of study candidates through each stage of the trial; 2) internal validity assessments comparing randomized study groups on correlates of mammography screening measured at three time points: the date of randomization, baseline survey completion, and the trial’s end; and 3) three external validity assessments—the first evaluating possible “cueing” effects of baseline and

interim surveys on mammography screening outcomes in our study population; the second comparing mammography screening trends over the 3.25-year study period in the study population, independent of intervention effects, with the corresponding trends in the US female population; and the third comparing characteristics of nonparticipants with those of participants remaining at the end of the trial. This approach can be adapted to future intervention trials in cancer prevention and control, regardless of whether they are efficacy trials of relatively limited scope or large-scale, population-based effectiveness trials.

Subjects and Methods

Study Design

Project HOME was a population-based randomized longitudinal intervention trial that was positioned closer to the effectiveness than the efficacy end of the continuum (4). As described in the accompanying article by Vernon et al. (13), Project HOME was designed to test hypotheses regarding intervention effects on mammography use during a 30-month follow-up period. Mammography rates in women who received the more personalized (tailored and targeted) intervention (group 1) and the less personalized (targeted) intervention (group 2) were compared with those in control group 3 (Figure 1) using intention-to-treat principles (18). An intention-to-treat analysis can improve both the internal and external validity of intervention trials by including the entire sample originally selected for the trial, regardless of participation or compliance with the intervention regimen (2).

Intervention trials can further address external validity by comparing nonparticipants with participants at the end of the trial on characteristics known to influence the study outcome (11). Substantive differences weaken the ability to generalize study findings to the target population of interest. For the conventional external validity assessments, Project HOME needed only three study groups. However, by adding control groups 4 and 5, for which the baseline survey measurements lagged 1 and 2 years, respectively, behind those of groups 1–3 (Figure 1), we designed Project HOME to address two other important external validity issues.

The first external validity issue arose because Project HOME was designed to test an intervention that incorporates personalized feedback tailored to individual participants' current knowledge, attitude, and behaviors relating to breast cancer screening, as described in Vernon et al. (13). To collect the information needed for tailoring, we administered preintervention surveys at baseline and year 1 follow-up (Figure 1). To examine the extent to which the baseline survey alone (independent of the intervention) served as a cue for women to obtain subsequent mammography screening, we compared control group 4 with control group 3. To determine whether the year 1 follow-up survey further influenced mammography screening (independent of the intervention), we compared control group 5 with control group 3 (Figure 1). Identifying any possible cueing effects of the baseline and year 1 follow-up surveys in this manner would strengthen the external validity of our final conclusions regarding intervention effects.

The other additional external validity issue relates to the longitudinal design of the trial and concern for possible changes (ie, secular trends) in the mammography screening behavior of our target population during the 3.25-year duration of the study, independent of any intervention influences. The causal factors underlying such population-wide temporal trends could be strong enough to modify (ie, augment or counteract) the trial's intervention effects. At the same time, if secular trends in the mammography screening behavior of our target population were found to coincide with trends reported for the general US female population, our results could be extrapolated beyond the target population of women veterans to the general US population.

Target Population

In the United States, opportunities to conduct effectiveness trials of cancer screening interventions are constrained by the health-care system available to the target population (19) and the variability in health-care providers' interpretation of the Health Insurance Portability and Accountability Act (20–22), both of which affect researchers' access to eligible study candidates and their medical records. We chose the occupational cohort of US women veterans because of its universal access to health-care services and the fact that it otherwise mirrors the US female population in demographic profile (23–26), geographic dispersion, mammography screening rates (27,28), and complex, real-world patterns of use of private- and government-sponsored health-care services (23,24,26). The Veterans Administration (VA) health-care system provides a “safety net” for uninsured and underinsured veterans (29), although more than 65% of US women veterans have never used the VA health-care system (26). Even regular VA users (<12% of all US women veterans) often alternate between VA and private-sector providers (30–32).

An estimated 39% of US women veterans were older than 50 years in 2002, similar to an estimated 41% of US females (33,34). The racial composition of US women veterans (71% non-Hispanic white, 18% African American, and 11% other) is also similar to that of the general US female population (71% non-Hispanic white, 13% African American, and 16% other). Finally, the proportion of women veterans with less than a college (bachelor's) degree (82%) is relatively close to that of the US female population (86%), although as a result of their service in the military more women veterans have received post-high school training (49%) than US females (27%) (33–35).

Source of Study Population (Sampling Frame)

The National Registry of Women Veterans (NRWV) was developed by one of the authors (D. J. del Junco, unpublished data, 1998) to promote health-related research of potential benefit to the entire population of US women veterans. The exact number of surviving women veterans since World War II is unknown due to historical limitations in data entry and processing and the lack of a coordinated system to track mortality in the veteran population. To create the NRWV, information on individual veterans had to be extracted from 14 different VA and Department of Defense (DoD) data sources. Each NRWV record identifies a unique woman veteran and contains the following basic set of variables: social security number (SSN), full name, date of birth, use of any VA health services between 1986 and 1997, and 14 indicators to identify which of the 14 VA and DoD source databases contains information about the veteran. Upon its completion in 1997, the NRWV identified 1.4 million of the 1.6 million women veterans that the VA estimated had separated from active military duty after January 1, 1942, and through December 31, 1997 (36). Separations of women from active military service since 1997 have been routinely updated in DoD and VA databases.

Study Sample Selection

Criteria for selection of potential study candidates from the NRWV included the following: a null or missing value for decedent status on the NRWV source databases, a valid SSN (needed to ascertain decedent status or current address), and an age of 52 years or older as of June 1, 2000. Despite controversies about the benefit of mammography screening in terms of mortality reduction (37–39), the US Preventive Services Task Force recommends that women aged 50–69 years be screened every 1–2 years with mammography (40), and the VA's mammography policy is to screen women 52–69 years of age every 1–2 years (41). An upper age limit was not imposed in our study because both medical judgment and recent evidence support screening older women unless otherwise indicated by health status (42).

To increase external validity (2), the data manager (SPC) used a computerized random number generator in Stata Statistical Software (Release 9; StataCorp LP, College Station, TX) to select the study sample from the subset of age-eligible NRWV records. To identify and exclude additional decedents and ineligible survivors with unknown age, we cross-linked the subset of age-eligible NRWV records with the updated VA Beneficiary Identification and Records Locator System (BIRLS) and Patient Treatment File (PTF) databases.

Additional deaths, males, and records with invalid SSNs or ineligible birthdates were identified and excluded from the study sample following cross-linkage with the Social Security Administration's Death Master File (43) and databases maintained by Experian. A comprehensive search strategy was developed to acquire and maintain up-to-date contact information on survivors in the study sample through databases maintained by Experian, the Internal Revenue Service, the National Change of Address database, and web-based public directories. Potential candidates were excluded if they lacked a complete, deliverable mailing address or if their only available address was outside the United States or Puerto Rico.

The study was conducted through The University of Texas Health Sciences Center–Houston School of Public Health and the VA Medical Center in Durham, North Carolina. The study protocol was approved by the institutional review boards of both institutions.

Participant Recruitment and Retention

To maximize our ability to systematically monitor study procedures, an in-house research team conducted all participant recruitment and retention activities. To increase the success of these procedures, we followed methods recommended by Dillman (44) and Aday (45). For example, we included small tokens of appreciation (eg, bumper sticker and jar and letter openers that were embossed with a Project HOME study logo and toll-free telephone number) in the first mailing of all study surveys. Nonrespondents to the first mailing of any study survey were sent a reminder postcard after 3 weeks. Nonrespondents to reminder postcards were sent a second survey after 7 weeks. After 11 weeks, if a telephone number was available, we called nonrespondents up to six times to give them an opportunity to provide information over the telephone. To maximize the success of our attempts at telephone contact, we followed a contingency-based protocol of two daytime calls, two evening calls, and two weekend calls. At the study's end, nonrespondents to the follow-up telephone calls in groups 1–3 were given an opportunity to win a \$500 lottery if they returned a completed abbreviated follow-up survey.

Because the NRWV is a historical database, we expected that only a portion of the records selected for Project HOME would ultimately fulfill all the eligibility criteria: a deliverable address; female sex; age 52 years or older as of June 1, 2000; past but not currently active service in the US military; no previous diagnosis of breast cancer; and no physical or mental inability to participate in the trial. To identify additional ineligible records after linkages with the available electronic databases, we mailed candidates a one-page eligibility survey along with a letter introducing the study. The introductory letter (and all subsequent study mailings) explained that participation in the trial was voluntary and involved responding to periodic mailed surveys and receiving educational materials, that all study information would be confidential, and that only study investigators would have access to the data. The mailed eligibility survey offered recipients the opportunity to report study ineligibility or decline study participation. Recipients who were eligible for the trial were asked to update their name and address information and to provide a telephone number for future contact if necessary. The eligibility survey was intended to maximize study group equality by enabling the exclusion of as many ineligible and uninterested individuals as possible before random assignment to the study groups.

To achieve the estimated required sample size of at least 3000 eligible participants (600 in each of the five groups) (13), we originally randomly sampled 12 000 records on September 4, 2000, to allow for nonresponse and ineligibility among potential candidates. Higher than expected levels of nonresponse ($n = 5582$) and ineligibility ($n = 4133$) following the initial mailing of the eligibility survey led us to draw a second random sample of 11 000 NRWV records on June 1, 2001. Although doing so was not part of the original study protocol, staggering the two sampling rounds 9 months apart provided a unique opportunity to evaluate the reliability of our sample selection, randomization, and recruitment procedures across time. Follow-up ended December 17, 2003, for sampling round 1 and October 1, 2004, for sampling round 2.

About 2 months after each round of eligibility surveys was mailed, the study's data manager conducted all random allocation procedures using Stata Statistical Software without knowledge of study candidates' NRWV characteristics or contact information. Random assignment to study groups 1–5 was stratified by sampling round (first or second) and candidates' status on the eligibility survey (self-reported eligible vs nonrespondent with unknown eligibility). The latter stratification enabled nonrespondents to join the study, report ineligibility, or refuse participation at any time during the 3.25-year study.

Within 1 month of random assignment, women in groups 1–3 were mailed a baseline survey. On completion and return of the baseline survey or 3 months after the baseline survey mailing date, whichever occurred first, we mailed women in groups 1 and 2 printed intervention materials that were targeted to the interests of women veterans and that encouraged them to be screened annually for breast cancer according to the American Cancer Society recommendations in effect at the time (46). As described in Vernon et al. (13), we also sent group 1 personalized letters and bookmarks containing expert (tailored) feedback (Figure 1). Approximately 1 year after mailing the baseline survey, we mailed the year 1 follow-up survey to the women in groups 1–3, and the 1-year–delayed baseline survey to the women in group 4. Within 2 months, we mailed the second intervention cycle to the women in groups 1 and 2. Approximately 1 year after mailing the year 1 follow-up survey, we mailed the final (year 2) follow-up survey to the women in groups 1–3 and the 2-year–delayed baseline survey to the women in group 5 (Figure 1).

Throughout the study, we ascertained deaths among potential study candidates from next of kin, local postmasters, and internet resources, including web-based genealogy death indexes and Experian databases. At the end of the study, we searched for any remaining but previously unreported decedents in the updated VA and Social Security Administration databases (43) and in the National Death Index (47).

Measurement of Variables

The baseline survey included questions on demographics, military experience, prior history of mammography screening, and other psychosocial factors as described in Vernon et al. (13). The year 1 and 2 follow-up surveys focused on mammography screening. Questions regarding ineligibility for Project HOME were included on all surveys.

For the internal validity assessments and the external validity assessment in which nonparticipants were compared with participants, we used available data on study candidate status (ie, survey respondent, nonrespondent, or postrandomization refusal to participate) along with the following factors associated with mammography screening (48–53): age (52–64 or ≥ 65 years, with the cut point based on Medicare eligibility), race/ethnicity (African American, Hispanic, non-Hispanic white, or unknown), educational level (high school or less, some college, college graduate or higher, or unknown), and the following surrogates for factors that are associated with mammography screening: current home state (CA, FL, TX, PA, NY, and all others), because regional differences in mammography screening rates across the United

States reflect population variation in income and access to care; self-report of past use of VA health-care services (yes or no) or source database of NRWV record = VA (VA PTF or Outpatient Clinic File [OPC], DoD, or VA BIRLS), because users of VA health-care services are known to have lower income and higher morbidity than other women veterans (30–32); sampling round (one or two), because study candidates recruited into the study at different time periods may experience different external influences on mammography screening behavior; and survey response mode (mail or phone), because the mammography screening behavior of participants who do not respond to mailed surveys but later respond to a telephone follow-up may differ from that of mailed survey respondents.

We ascertained mammography status according to self-report on the baseline and postintervention year 1 and year 2 follow-up surveys. On each survey, respondents were asked to report the month and year of their two most recent mammograms. Although trials using intention-to-treat analyses ideally measure the study outcome of everyone in the study sample regardless of study participation (eg, by using an existing database of medical records) (4, 18), this is not typically an option for population-based trials that are conducted outside organized systems of health care. In Project HOME, we were able to search the VA's centralized medical records databases (ie, the PTF or OPC) between October 1, 1996, and September 30, 2005, as an independent source of mammography ascertainment for the subset of women who had ever used VA health-care services. For these analyses, study women were identified as an "ever VA user" if they reported having VA health-care insurance or receiving their last mammogram at a VA-sponsored facility on their baseline survey, if either the VA PTF or OPC database was a source of their original NRWV record, or if a VA record documented their receipt of one or more VA-sponsored mammograms within the 30 months before the date their baseline survey was mailed.

Statistical Analyses

Internal Validity Assessment—We tested for any inequalities in the distributions of covariates across study groups using univariate chi-square tests at three time points: the date of randomization, the date of baseline survey completion, and the date of completion of the final follow-up survey. For all time points, we compared the study groups on study candidate status, sampling round, and age group. At randomization, we also compared groups on current home state and NRWV source database. At baseline survey completion, we compared groups on survey response mode, race/ethnicity, level of education, and ever VA user status. We also tested between-group differences in mean age at baseline using one-way analysis of variance. At the end of the study, we compared groups 1–3 on the following additional variables: pattern of response to all three postrandomization surveys, final follow-up survey response mode (mail, telephone, or abbreviated survey), and ever VA user status.

External Validity Assessment: Evaluation of Possible Cueing Effects of the Baseline and Year 1 Follow-up Surveys—To evaluate the possible cueing effects of the baseline survey, we compared group 4 with group 3 (referent) in terms of self-reported mammography experience during the time interval between the dates that the baseline and year 1 follow-up surveys were mailed to group 3 (Figure 1). We defined coverage as the report of at least one mammogram within the year 1 time interval.

To evaluate the possibility of further cueing effects associated with the year 1 follow-up survey, we compared group 5 with group 3 (referent) in terms of self-reported mammography experience during the time interval between the dates that the baseline and year 2 follow-up surveys were mailed to group 3 (Figure 1). We defined compliance as the report of at least two mammograms 6–15 months apart within the combined total year 1 and year 2 time intervals. We extended the annual mammography interval for measuring compliance from 12 to 15

months to allow for scheduling and insurance constraints (54,55). We assessed the statistical significance of between-group differences with chi-square tests.

External Validity Assessment: Secular Trend Analyses—We examined secular trends in mammography screening across the 3.25-year study period, independent of intervention trial influences, by comparing all study groups (using group 3 as the referent) on recent mammography history within the 30 months before their respective dates (Figure 1) of baseline survey completion. We defined coverage as the report of at least one mammogram within the 30-month period preceding completion of each group’s baseline survey and compliance as the report of at least two mammograms 6–15 months apart within the same 30-month period.

We first examined secular trends in self-reported mammography screening for all groups using data from their respective baseline surveys. In a second analysis, we focused on the subset of ever VA users to examine secular trends in mammography screening as documented in computerized VA records, independent of baseline survey completion. Mammography status for ever VA users with no mammogram dates recorded in the VA databases was coded “0,” a “worst-case” method of imputation (56). Finally, we compared the self-reported baseline mammography screening rates in our study population with the corresponding rates reported for the general US female population (57,58). The statistical significance of differences between groups was assessed with chi-square tests.

External Validity Assessment: Comparison of Nonparticipants With Participants at Study End—We compared the characteristics of nonparticipants with those of study participants retained through the end of the trial in groups 1–3. Participants were study candidates who provided complete mammography information on either their year 1 or year 2 follow-up surveys. Nonparticipants in groups 1–3 were first divided into two general categories, “presumed eligible” and “reported ineligible.” Presumed eligible candidates include those who refused to participate and those who never responded to any study survey, that is, “never respondents.” Reported ineligible candidates include decedents, nonveterans, males, candidates under 52 years of age, those with no forwarding address, those with a history of breast cancer, those on active military duty, and those physically or mentally unable to participate. To examine between-group differences in the proportion of nonparticipants relative to participants after adjustment for covariates, we compared each subtype of presumed eligible non participants with participants using multiple logistic regression models that included the following covariates: age group, sampling round, and NRWV source database. All data analyses were conducted by the data manager using Stata Statistical Software.

Results

Flow Diagrams: Passage of Study Candidates Through Each Stage of the Trial

Of the total 1.4 million NRWV records, records for 83 292 individuals were excluded due to missing or invalid SSN (Figure 2) and represent older, pre–Vietnam Conflict veterans with no record of VA benefits. A total of 177 068 records formed the initial pool of potential age-eligible study candidates from which we drew a total random sample of 23 000 records (Figure 2) in two sampling rounds (ie, 12 000 in round 1 and 11 000 in round 2). Linkages with the updated VA databases led to exclusions due to errors in SSNs, ineligible birthdates, and death; these exclusions reduced the final candidate pool to 160 719 and the final random sample to 21 449 (11 448 from round 1 and 10 001 from round 2). As expected, due to deaths among older individuals during the 9 months between the two sampling rounds, average birth years differed (median = 1936 for sampling round 1, median = 1937 for sampling round 2, $P = .005$). Also as expected, a greater number of regular users of VA health services had died by the time

we selected potential study candidates in sampling round 2. Thus, the source databases for sampling round 1 records were more often VA OPC or PTF (27%) than the source databases for sampling round 2 records (24%, $P = .001$). Sampling rounds 1 and 2 did not differ on state of residence ($P = .65$), and sampling round did not predict either a refusal to participate (odds ratio [OR] = 0.97, 95% confidence interval [CI] = 0.86 to 1.09, referent = round 2) or a report of ineligibility for any reason (OR = 1.10, 95% CI = 0.99 to 1.22, referent = round 2) in response to the eligibility survey. Therefore, we combined the two sampling rounds to simplify the flow diagram in Figure 2.

After excluding additional ineligible individuals due to death, male sex, or ineligible birthdate (Figure 2), we acquired addresses and sent eligibility surveys to a total of 16 341 potential study candidates. Response to the eligibility survey reduced the list of study candidates to a total of 4048 eligible candidates and 10 366 nonrespondents with unknown eligibility status. Potential study candidates were randomly assigned in roughly equal numbers to one of the five study groups (Figure 2). Before mailing the baseline surveys, we excluded an additional 321 ineligible participants who were identified from eligibility surveys that had been received after the randomization cutoff date, which left a total of 10 045 nonrespondents with unknown eligibility status.

The 14093 study candidates in groups 1–5 were tracked from the date of their baseline survey mailings to the end of the study (Figure 3). By the trial's end, groups 1–3 had a total of 3126 participants and 5311 nonparticipants for the intervention effectiveness analyses and for the third external validity assessment. Of the nonparticipants in groups 1–3, 2157 were presumed eligible and 3154 were reported ineligible (Figure 3). Of those presumed eligible in groups 1–3, 447 never responded and 1710 refused (Figure 3). Whereas the parallel groups 1–3 had similar numbers reported ineligible within each subcategory, the marked differences in the numbers reported ineligible for groups 4 and 5 (shaded in Figure 3) reflect the time lags and fewer study contacts imposed by the study's design (Figure 1). By the study's end, we were able to establish study eligibility for 21 340 (92.8%) of the 23 000 records drawn from the NRWV.

Internal Validity Assessment

Study Groups 1–5 at Randomization—At the time of randomization, the five study groups did not differ with respect to any of the covariates tested (Table 1). As expected, the stratification variables (ie, study candidate status and sampling round) showed the least between-group variation. Study candidates' addresses spanned all 50 states, the District of Columbia, and Puerto Rico. Also as expected for this older population (59), the states of California, Florida, Texas, New York, and Pennsylvania accounted for 35% of the study population. Study candidates ranged in age from 52 to 100 years (mean = 62.38 years, median = 58 years), with 95% of the study population aged 80 years or younger.

Study Groups 1–5 at Respective Baseline Survey Completion—Despite the time lags in baseline survey mailings (Figure 1), respondents across the five study groups were similar with respect to race/ethnicity, education, ever use of VA health-care services, sampling round, and age group (Table 2). Mean age at the time of the baseline survey differed in the manner expected across the five groups (63.1, 62.9, 63.1, 64.0, and 64.7 years, respectively; $P < .001$). Compared with groups 1–3, groups 4 and 5 were less likely to respond to the baseline survey or to follow-up telephone calls (Table 2).

Study Groups 1–3 at Trial's End—The patterns of study candidates' completion of the baseline and postintervention follow-up surveys did not differ across the three groups assigned to the intervention trial (Table 3). At study's end, there were no between-group differences in

participants' follow-up survey response mode, ever use of VA health-care services, sampling round, or age group.

External Validity Assessment

Evaluation of Possible Cueing Effects of the Baseline and Year 1 Follow-up Surveys—At the time of the year 1 follow-up survey, 634 of the 900 group 3 respondents (70%) reported receiving at least one mammogram (ie, mammography coverage) since the date their baseline survey was mailed and 764 of the 1091 group 4 respondents (70%; $P = .84$) reported receiving at least one mammogram in the year before their baseline survey. At the time of the year 2 follow-up survey, 487 of the 802 group 3 respondents (61%) reported receiving at least two mammograms 6–15 months apart (ie, mammography compliance) since the date their baseline survey was mailed compared with 599 of the 1020 group 5 respondents (59%, $P = .388$) in the 2 years before their baseline survey.

Secular Trend Analyses in Study Groups 1–5—Compared with the referent, group 3, self-reported mammography coverage and compliance at baseline survey did not differ at a statistically significant level for any of the other groups (Table 4). However, coverage in group 5 (82.3%) was lower than that of any other group, and when compared with that in groups 1–4 combined (85.1%), the difference was statistically significant ($P = .024$).

In the subgroup of those who had ever used VA health-care services, which included eligible study candidates regardless of survey completion, group 5 was the only group that differed from group 3 at a statistically significant level on mammography coverage based on VA records (Table 4). Group 5 also had lower rates of both mammography coverage ($P = .002$) and mammography compliance ($P = .077$) based on VA records than groups 1–4 combined. Overall, the data in Table 4 suggest a drop in mammography screening rates in the final year of the 3.25-year study. A similar decrease has been reported in the US female population between the years 2000 and 2005 (57,58). Of interest, among those who had ever used VA health-care services, the mammography screening rates based on VA records for nonrespondents to the baseline survey ($n = 1489$, 11.6% coverage and 5.5% compliance) were substantially lower than the corresponding rates for baseline survey respondents ($n = 2187$, 28.1% coverage, $P < .001$, and 14.4% compliance, $P < .001$). In other words, when mammography screening was measured by more objective medical records, respondents to the baseline survey had higher mammography screening rates than nonrespondents.

In the 2000 National Health Interview Survey (60), 79% of US women aged 50–64 years reported at least one mammogram within the preceding 24 months (ie, coverage); among respondents to the Project HOME baseline survey the rate was 80%. Similarly, the average rate of repeat mammography reported in a recent systematic review by Clark et al. (54) for women aged 50 years and older who were surveyed between 1995 and 2001, 57%, was very close to the 58% rate of self-reported compliance at baseline (ie, two mammograms 6–15 months apart) in Project HOME.

Comparison of Nonparticipants With Participants in Groups 1–3 at Study's End—By the end of the trial, the distribution of the two subcategories of presumed eligible nonparticipants (Figure 3) was similar across study groups 1–3 ($P = .67$). Even after adjustment for available covariates using multiple logistic regression analysis (Table 5), there were no statistically significant between-group differences in the rates of either refusal or never response. There were no statistically significant interactions between study group and the other covariates, and non-statistically significant interaction terms were dropped from the final logistic models in Table 5. Regardless of study group, candidates identified from the VA OPC and PTF databases were more likely to refuse or to never respond than those identified from

the VA BIRLS databases, and those in sampling round 2 were more likely to never respond than those in sampling round 1 (Table 5). Compared with older women, study candidates 52–64 years of age were less likely to refuse participation but more likely to never respond.

Discussion

We demonstrated the feasibility of a population-based health promotion trial and an approach to assess both internal and external validity. The flow diagrams that tracked the passage of study candidates through each stage of the trial and the results from our internal validity assessments across the trial's three key time points provided evidence of Project HOME's internal validity. The results from our three different assessments of external validity provided evidence of Project HOME's external validity. To date, most trials of health promotion interventions (1) address internal validity to some extent but provide limited, if any, evidence of external validity. Even among intervention trials using random allocation to increase internal validity, few have assessed the potential threats to internal (as well as external) validity due to nonparticipation among randomized study candidates. Target populations are seldom clearly defined, and the extent to which the final group of study participants represents the target population is rarely evaluated.

There are challenges in conducting health promotion research using a historical database such as the NRWV as a research sampling frame. Although these challenges are well known among researchers who conduct long-term morbidity and mortality surveillance studies in large occupational and community cohorts (61–64), they are less familiar to health promotion researchers. The NRWV posed several potential limitations for Project HOME. For example, because of inconsistencies in the computerization of military and VA records before the 1990s, the records of some older women veterans who never applied for VA benefits are missing from the NRWV. Nevertheless, the NRWV can be considered a virtually exhaustive sampling frame for surviving older women veterans who have ever applied for VA benefits and a large sampling frame with no known selection bias for all other surviving older women veterans. Tracing an occupational cohort of women is particularly challenging due to surname changes and the tendency to erroneously substitute a spouse's SSN on official records (65–67). Despite these difficulties, we were able to establish study eligibility for 93% of the 23 000 potential candidates who were randomly selected from the age-eligible subset of the NRWV.

The five study groups were similar at randomization and at the time they completed their respective baseline surveys on all covariates tested except response rates to the baseline surveys. The lower response rates in groups 4 and 5 (for whom the baseline surveys were delayed 1 and 2 years, respectively, to facilitate external validity assessments) were additional study limitations. Extensive evidence from the survey research literature attests to the marked decrease over time in the response of study candidates to telephone surveys (68,69). Therefore, it is increasingly important for researchers to establish and maintain multiple modes of communication between study candidates and the research team (70,71). In contrast to the lower survey participation in groups 4 and 5 of Project HOME, study groups 1–3 remained similar with respect to their distributions on all measured covariates throughout all stages of the trial. In addition, mammography history, one of the most important predictors of subsequent mammography screening behavior (48–53), was similar in groups 1–3, whether ascertained by self-report or from VA records. Thus, the results from the intervention effectiveness analyses described in Vernon et al. (13) are likely to be internally valid.

In the first external validity assessment, receipt of the baseline and year 1 follow-up surveys appeared to have negligible influence on subsequent self-reported mammography screening behavior. Therefore, survey cueing effects would be unlikely to explain any statistically significant differences in postintervention mammography screening rates among groups 1–3.

Results from the second validity assessment regarding secular trends in the target population's mammography screening behavior, independent of the trial's influence, suggest a slight decrease between the years 2001 and 2004 regardless of the source of mammography information (self-report or VA record). Similar rates of mammography coverage and compliance in the US female population (54,60) and the decline reported for the US female population between 2000 and 2005 (57,58) lend additional support to the generalizability of our results. Nevertheless, Project HOME was limited in its ability to ascertain mammography use objectively because the majority of women veterans have never used VA health-care services and even regular users often alternate between VA and private health-care providers (30–32). In addition, VA records do not systematically document mammograms received outside the VA system. Restricting our ascertainment of mammography based on VA records to the subgroup of ever users of VA health-care services reduced the degree to which we underestimated true mammography screening rates in our intention-to-treat analyses. Randomization tends to balance the degree of underestimation across study groups, and this type of nondifferential misclassification (72) would be expected to produce slightly more (not less) conservative intervention effect estimates (ie, closer to the null).

An additional related limitation in Project HOME was the large number of presumably eligible nonparticipants (2157, vs 3126 eligible participants for groups 1–3). In our comparisons of presumed eligible nonparticipants with eligible participants, there were no statistically significant between-group differences from either the univariate or multivariable analyses that included all available covariates. However, the observed statistically significant differences between nonparticipants and participants on covariates independent of study group (eg, age, sampling round, NRWV source database) would bias absolute but not relative measures of intervention effects because the covariate distributions were balanced across the study groups. Similarly, restricting our outcome analyses to study participants could overestimate the true absolute between-group differences in the total study population if eligible nonparticipants had lower postintervention mammography rates. Indeed, in the subset of women who had ever used VA health-care services, mammography coverage and compliance rates based on VA records were lower in nonrespondents than in respondents to our baseline survey. Because we used intention-to-treat analyses and computed relative (not absolute) between-group differences (13), the trial's relative intervention effect estimates would be biased only if eligible nonparticipants differed materially from participants on important but unmeasured confounders (72). In the report by Vernon et al. (13), we were able to examine the net effect of such residual confounding in the subgroup of women who had ever used VA health-care services by comparing intervention effect estimates from analyses based on postintervention mammograms reported in VA records with those based on self-reported mammograms. The differences were consistently negligible, supporting the external validity of our relative measures of intervention effects.

In the absence of a universal health-care system, measuring the effectiveness of cancer prevention and control strategies is challenging because it requires the study of large, well-defined, representative populations with access to health care (19). Clinical cancer research in the United States has benefited from the existence of research networks like the Community Clinical Oncology Programs (73), but until recently such networks have not been available for cancer prevention research. The National Cancer Institute–funded Cancer Research Network, which has been in existence since 1999, provides a multicenter collaborative setting in which to pursue more generalizable cancer-related health promotion research among persons with access to health care (74). The occupational cohort of veterans provides a nationally representative target population capable of supporting highly generalizable research in cancer prevention and control because all veterans are ensured access to health care through the VA's safety net (29), and objective measures of study outcomes are available for a large subset of the target population through the VA's electronic databases. Moreover, results from our

internal and external validity assessments suggest that Project HOME's findings regarding intervention effectiveness can be generalized, beyond the female veteran cohort to the entire US population of women aged 52 years and older.

To our knowledge, this is the first study in the cancer screening intervention literature to systematically examine and provide evidence for both internal and external validity. Although standards of practice relating to external validity are in an earlier stage of development than those for internal validity (4,11,14,17), meeting the National Institutes of Health objective to develop effective interventions that can be widely disseminated (5) will require evidence of external as well as internal validity. Successful intervention development and dissemination will depend as much on the validity of trial findings as on the innovativeness of the interventions. The study design and data analysis strategies that formed the basis of the systematic validity assessment described here can be readily adapted to future intervention trials.

CONTEXT AND CAVEATS

Prior knowledge

Trials to assess interventions to change behavior often evaluate internal validity or the efficacy of the intervention under optimal conditions. However, they are rarely set up to evaluate external validity, that is, the intervention's likely effectiveness in the broad population.

Study design

A nationwide, population-based trial of US women veterans to evaluate an intervention to promote regular mammography, Project Healthy Outlook on the Mammography Experience, was designed to address both internal and external validity. Internal validity was assessed by comparing intervention groups on participation and factors associated with mammography screening; external validity was assessed by establishing additional control groups to evaluate survey "cueing" effects and general population trends and by comparing characteristics of nonparticipants with those of participants.

Contributions

The intervention groups were similar, indicating that the trial is likely to be internally valid. There was no evidence that completing the baseline or interim follow-up surveys by themselves led to an increase in mammography. The external validity assessment revealed a decline in mammography over time similar to that seen in the broader US population. Finally, there were no statistically significant between-group differences comparing nonparticipants with participants.

Implications

The trial is likely to be both internally and externally valid. The approach used by the authors is adaptable to studies of other behavior change interventions.

Limitations

Many women veterans obtain health care outside the Veterans Administration system, so complete ascertainment of mammography in nonparticipants was not possible. This type of under-ascertainment would be expected to reduce only slightly the magnitude of the study's intention-to-treat estimates of intervention effectiveness.

Acknowledgments

Funding

J Natl Cancer Inst. Author manuscript; available in PMC 2010 March 29.

National Institutes of Health (RO1 NCI CA76330 to L.A.B., W.C., D.J.d.J., D.R.L., W.R., J.A.T., S.W.V., S.P.C.; RO3 NCI CA 103512 to D.J.d.J., L.S.S., S.W.V.).

The authors take sole responsibility for the study design, data collection and analyses, interpretation of the data, and the preparation of the manuscript. The authors are indebted to the US women veterans who participated in the study and to the members of the external advisory committee who reviewed study materials and assisted in various other ways throughout the study. We thank Amy Jo Harzke for her comments on an earlier draft of the manuscript.

References

1. Hallfors D, Cho H, Sanchez V, Khatapoush S, Kim HM, Bauer D. Efficacy vs effectiveness trial results of an indicated “model” substance abuse program: implications for public health. *Am J Public Health* 2006;96(12):2254–2259. [PubMed: 16809591]
2. Last, JM. *A Dictionary of Epidemiology*. 4th ed.. New York, NY: Oxford University Press; 2001.
3. Flay BR. Efficacy and effectiveness trials (and other phases of research) in the development of health promotion. *Prev Med* 1986;15:451–474. [PubMed: 3534875]
4. Gartlehner, G.; Hansen, RA.; Nissman, D.; Lohr, KN.; Carey, TS. Criteria for distinguishing effectiveness from efficacy trials in systematic reviews. Rockville, MD: Agency for Healthcare Research and Quality, US Department of Health and Human Services; Prepared by the RTI-International-University of North Carolina Evidence-based Practice Center under Contract No. 290-02-0016. AHRQ Publication No. 06-0046
5. Department of Health and Human Services. Dissemination and Implementation Research in Health (R01): PAR-06-039. 2006 [Accessed July 25, 2007]. <http://grants.nih.gov/grants/guide/pa-files/PAR-06-039.html>
6. Glasgow RE, Marcus AC, Bull SS, Wilson KM. Disseminating effective cancer screening interventions. *Cancer* 2004;101(5 suppl):1239–1250. [PubMed: 15316911]
7. Kerner JF, Rimer B. Introduction to the special section on dissemination: dissemination research and research dissemination: how can we close the gap? *Health Psychol* 2005;24(5):443–446. [PubMed: 16162037]
8. Glasgow RE, Davidson KW, Dobkin PL, Ockene JK, Spring B. Practical behavioral trials to advance evidence-based behavioral medicine. *Ann Behav Med* 2006;31(1):5–13. [PubMed: 16472033]
9. Harris R. Effectiveness: the next question for breast cancer screening. *J Natl Cancer Inst* 2005;97(14):1021–1023. [PubMed: 16030293]
10. Meissner HI, Vernon SW, Rimer BK, et al. The future of research that promotes cancer screening. *Cancer* 2004;101(5 suppl):1251–1259. [PubMed: 15316910]
11. Glasgow RE, Lichtenstein E, Marcus AC. Why do n’t we see more translation of health promotion research to practice? Rethinking the efficacy-to-effectiveness transition. *Am J Public Health* 2003;93:1261–1267. [PubMed: 12893608]
12. Drummond, MF.; Sculpher, MJ.; Torrance, GW.; O’Brien, BJ.; Stoddart, GL. *Methods for the Economic Evaluation of Health Care Programmes*. 3rd ed.. Oxford, England: Oxford University Press; 2005.
13. Vernon SW, del Junco DJ, Tiro JA, et al. Promoting regular mammography screening II. Results from a randomized controlled trial in US women veterans. *J Natl Cancer Inst* 2008;100(5):347–358. [PubMed: 18314474]
14. Moher D, Schulz KF, Altman D. for the CONSORT Group. The CONSORT statement: revised recommendations for improving the quality of reports of parallel-group randomized trials. *JAMA* 2001;285(15):1987–1991. [PubMed: 11308435]
15. AAPOR Council. Code of professional ethics and practices. 2007 [Accessed February 1, 2008]. <http://www.aapor.org/aaporcodeofethics>
16. The American Association for Public Opinion Research. Standard Definitions: Final Dispositions of Case Codes and Outcome Rates for Surveys. 4th ed.2006 [Accessed February 1, 2008]. p. 1-42.<http://www.aapor.org/uploads/standarddefs4.pdf>
17. Davidson KW, Goldstein MG, Kaplan RM, et al. Evidence-based behavioral medicine: what is it and how do we achieve it? *Ann Behav Med* 2003;26(3):161–171. [PubMed: 14644692]

18. Lachin JM. Statistical considerations in the intent-to-treat principle. *Control Clin Trials* 2000;21(3): 167–189. [PubMed: 10822117]
19. Miles A, Cockburn J, Smith RA, Wardle J. A perspective from countries using organized screening programs. *Cancer* 2004;101(5 suppl):1201–1213. [PubMed: 15316915]
20. Gostin LO. National health information policy—regulations under the health insurance portability and accountability act. *JAMA* 2001;285(23):3015–3021. [PubMed: 11410101]
21. Gunn PP, Fremont AM, Bottrell M, Shugarman LR, Galegher J, Bikson T. The health insurance portability and accountability act privacy rule. A practical guide for researchers. *Med Care* 2004;42(4):321–327. [PubMed: 15076808]
22. Hodge JG, Gostin LO, Jacobson PD. Legal issues concerning electronic health information. *JAMA* 1999;282(15):1466–1471. [PubMed: 10535438]
23. Boyle, JM. Survey of Female Veterans: A Study of the Needs, Attitudes and Experiences of Women Veterans. [Study Conducted for the Veterans Administration]. New York, NY: Harris (Louis) and Associates, Inc; 1985. Louis Harris and Associates. August Report 843002
24. National Center for Veterans Analysis and Statistics. National Survey of Veterans. Washington, DC: Department of Veterans Affairs; 1995. April Report NSV9503 (depot stock no. P92493)
25. Estimates and Projections of the Veteran population, 1990–2030, Vetpop 2001. Washington, DC: Department of Veterans Affairs, Office of the Actuary; [Accessed August 1, 2007]. <http://www1.va.gov/vetdata/docs/51.xls>
26. 2001 National Survey of Veterans (NSV): Final Report. 2004 Mar 26 [Accessed August 1, 2007]. p. 1-139. http://www1.va.gov/vetdata/docs/survey_final.htm.
27. Hynes DM, Bastian LA, Rimer BK, Sloane R, Feussner JR. Predictors of mammography use among women veterans. *J Womens Health* 1998;7(2):239–247. [PubMed: 9555689]
28. Goldzweig CL, Parkerton PH, Washington DL, Lanto AB, Yano EM. Primary care practice and facility quality orientation: influence on breast and cervical cancer screening rates. *Am J Manag Care* 2004;10(4):265–272. [PubMed: 15124503]
29. Simpson SA. Safety net lessons from the veterans health administration. *Am J Public Health* 2006;96(6):956. [PubMed: 16670213]
30. Washington DL, Yano EM, Simon B, Sun S. To use or not to use: what influences why women veterans choose VA health care. *J Gen Intern Med* 2006;21:S11–S18. [PubMed: 16637939]
31. Yano EM, Washington DL, Goldzweig CL, Caffrey C, Turner RN. The organization and delivery of women's health care in Department of Veterans Affairs medical center. *Womens Health Issues* 2003;13(2):55–61. [PubMed: 12732441]
32. Nelson KM, Starkebaum GA, Reiber GE. Veterans using and uninsured veterans not using Veterans Affairs (VA) health care. *Public Health Rep* 2007;122(1):93–100. [PubMed: 17236614]
33. Department of Veterans Affairs. Adjusted estimates and projections of the female veteran population by race/ethnicity, 2000–2033. Washington, DC: Department of Veterans Affairs, Office of the Actuary; 2007. Table of female veterans by race/ethnicity, 2002
34. U.S. Census Bureau. Population Estimates Program. Washington, DC: Population Division; 2006.
35. Department of Veterans Affairs. VetPop2004. Adjusted estimates and projections of the female veteran population by state, 2000–2033. [Accessed February 1, 2008]. Table 1L: veterans by state, age group, period, gender 2000–2033. <http://www1.va.gov/vetdata/docs/VP2004B.htm>
36. Richardson, C.; Waldrop, J. Veterans: 2000. Census 2000 Brief. 2003 [Accessed May 25, 2007]. <http://www.census.gov/prod/2003pubs/c2kbr-22.pdf>
37. Humphrey LL, Helfand M, Chan BKS, Woolf SH. Breast cancer screening: a summary of the evidence for the U.S. Preventive Services Task Force. *Ann Intern Med* 2002;137(5 pt 1):347–360. [PubMed: 12204020]
38. Gotzsche PC, Olsen O. Is screening for breast cancer with mammography justifiable? *Lancet* 2000;355:129–134. [PubMed: 10675181]
39. Goodman SN. The mammography dilemma: a crisis for evidence-based medicine? *Ann Intern Med* 2002;137(5 pt 1):363–365. [PubMed: 12204023]

40. U.S. Preventive Services Task Force. Screening for breast cancer: re commendations and rationale. Rockville, MD: Agency for Healthcare Research and Quality; [Accessed May 23, 2007]. <http://www.ahrq.gov/clinic/3rduspstf/breastcancer/brcanrr.htm>
41. Department of Veterans Affairs. Preventive services included within National Performance Measures FY 2007. National Center for Health Promotion & Disease Prevention (NCP). 2007 May [Accessed May 23, 2007]. <http://www.prevention.va.gov/docs/FY2007PreventiveServicesPMFY07.doc>
42. Galit W, Green MS, Lital KB. Routine screening mammography in women older than 74 years: a review of the available data. *Maturitas* 2007;57(2):109–119. [PubMed: 17336004]
43. National Technical Information Service (NTIS). Social Security Administration Death Master File. Springfield, VA: US Department of Commerce Technology Administration; 2007.
44. Dillman, DA. Mail and Telephone Surveys: The Total Design Method. New York, NY: John Wiley & Sons; 1978.
45. Aday, LA. Designing and Conducting Health Surveys: A Comprehensive Guide. 2nd ed.. San Francisco, CA: Jossey-Bass; 1996.
46. Greenlee RT, Murray T, Bolden S, Wingo PA. Cancer statistics, 2000. *CA Cancer J Clin* 2000;50(1): 7–33. [PubMed: 10735013]
47. National Death Index. 2007 [Accessed July 31, 2007]. <http://www.cdc.gov/nchs/ndi.htm>
48. Selvin E, Brett KM. Breast and cervical cancer screening: sociodemographic predictors among White, Black, and Hispanic women. *Am J Public Health* 2003;93(4):618–623. [PubMed: 12660207]
49. Rakowski W, Breen NL, Meissner HI, et al. Prevalence and correlates of repeat mammography among women aged 55–79 in the year 2000 National Health Interview Survey. *Prev Med* 2004;39(1):1–10. [PubMed: 15207980]
50. Rakowski W, Meissner H, Vernon SW, Breen N, Rimer B, Clark MA. Correlates of repeat and recent mammography for women ages 45 to 75 in the 2002 to 2003 health information national trends surveys (HINTS 2003). *Cancer Epidemiol Biomarkers Prev* 2006;15(11):2093–2101. [PubMed: 17119033]
51. Champion VL, Maraj M, Hui SL, et al. Comparison of tailored interventions to increase mammography screening in nonadherent older women. *Prev Med* 2003;36(2):150–158. [PubMed: 12590989]
52. Hiatt RA, Klabunde CN, Breen NL, Swan J, Ballard-Barbash R. Cancer screening practices from National Health Interview Surveys: past present, and future. *J Natl Cancer Inst* 2002;94(24):1837–1846. [PubMed: 12488477]
53. Sohl SJ, Moyer A. Tailored interventions to promote mammography screening: a meta-analysis review. *Prev Med* 2007;45(4):252–261. [PubMed: 17643481]
54. Clark MA, Rakowski W, Bonacore LB. Repeat mammography: prevalence estimates and considerations for assessment. *Ann Behav Med* 2003;26(3):201–211. [PubMed: 14644696]
55. Partin MR, Casey-Paal AL, Slater JS, Korn JE. Measuring mammography compliance: lessons learned from a survival analysis of screening behavior. *Cancer Epidemiol Biomarkers Prev* 1998;7(8):681–687. [PubMed: 9718220]
56. Wood AM, White IR, Thompson SG. Are missing outcome data adequately handled? A review of published randomized controlled trials in major medical journals. *Clin Trials* 2004;1(4):368–376. [PubMed: 16279275]
57. Breen N, Cronin KA, Meissner HI, et al. Reported drop in mammography: is this cause for concern. *Cancer* 2007;109(12):2405–2409. [PubMed: 17503429]
58. Chagpar AB, McMasters KM. Trends in mammography and clinical breast examination: a population-based study. *J Surg Res* 2007;140(2):214–219. [PubMed: 17418862]
59. U.S. Census Bureau. Statistical brief: sixty-five plus in the United States. 2001 Jul 27 [Accessed June 25, 2007]. <http://www.census.gov/population/socdemo/statbriefs/agebrief.html>
60. Health, United States, 2006. With Chartbook on Trends in the Health of Americans. Hyattsville, MD: National Center for Health Statistics; 2006 [Accessed July 31, 1907]. p. 313 <http://www.cdc.gov/nchs/products/pubs/pubd/hs/women.htm#ambulatory>
61. Schall LC, Buchanich JM, Marsh GM, Bittner GM. Utilizing multiple vital status tracing services optimizes mortality follow-up in large cohort studies. *Ann Epidemiol* 2001;11(5):292–296. [PubMed: 11399442]

62. Hunt JR, White E. Retaining and tracking cohort study members. *Epidemiol Rev* 1998;20(1):57–70. [PubMed: 9762509]
63. Harvey BJ, Wilkins AL, Hawker GA, et al. Using publicly available directories to trace survey nonresponders and calculate adjusted response rates. *Am J Epidemiol* 2003;158(10):1007–1011. [PubMed: 14607809]
64. Pirzada A, Yan LL, Garside DB, Schiffer L, Dyer AR, Daviglius ML. Response rates to a questionnaire 26 years after baseline examination with minimal interim participant contact and baseline differences between respondents and nonrespondents. *Am J Epidemiol* 2004;159(1):94–101. [PubMed: 14693664]
65. Hill ME, Preston SH, Rosenwaike I. Age reporting among white Americans aged 85+: results of a record linkage study. *Demography* 2000;37(2):175–186. [PubMed: 10836175]
66. Niedhammer I, Saurel-Cubizolles MJ, Piciotti M, Bonenfant S. How is sex considered in recent epidemiological publications on occupational risks? *Occup Environ Med* 2000;57(8):521–527. [PubMed: 10896958]
67. Zahm S, Pottner L, Lewis D, Ward M, White D. Inclusion of women and minorities in occupational cancer epidemiologic research. *J Occup Med* 1994;36(8):842–847. [PubMed: 7807263]
68. Curtin R, Presser S, Singer E. Changes in telephone survey nonresponse over the past quarter century. *Public Opin Q* 2005;69(1):87–98.
69. Tuckel P, O'Neill H. The vanishing respondent in telephone surveys. *J Advert Res* 2002;42(5):26–48.
70. Leeuw ED. To mix or not to mix data collection modes in surveys. *J Off Stat* 2005;21(2):233–255.
71. Galea S, Tracy M. Participation rates in epidemiologic studies. *Ann Epidemiol* 2007;17(9):643–653. [PubMed: 17553702]
72. Rothman, KJ.; Greenland, S. *Modern Epidemiology*. 2nd ed.. Philadelphia, PA: Lippincott-Raven; 1998.
73. National Cancer Institute Division of Cancer Prevention (DCP). Community Clinical Oncology Program (CCOP). 2007 [Accessed May 23, 2007]. <http://prevention.cancer.gov/programs-resources/programs/ccop/about/facts>
74. Wagner EH, Greene SM, Hart TS, et al. Building a research consortium of large health systems: the cancer research network. *J Natl Cancer Inst Monographs* 2005;35:3–11. [PubMed: 16287880]

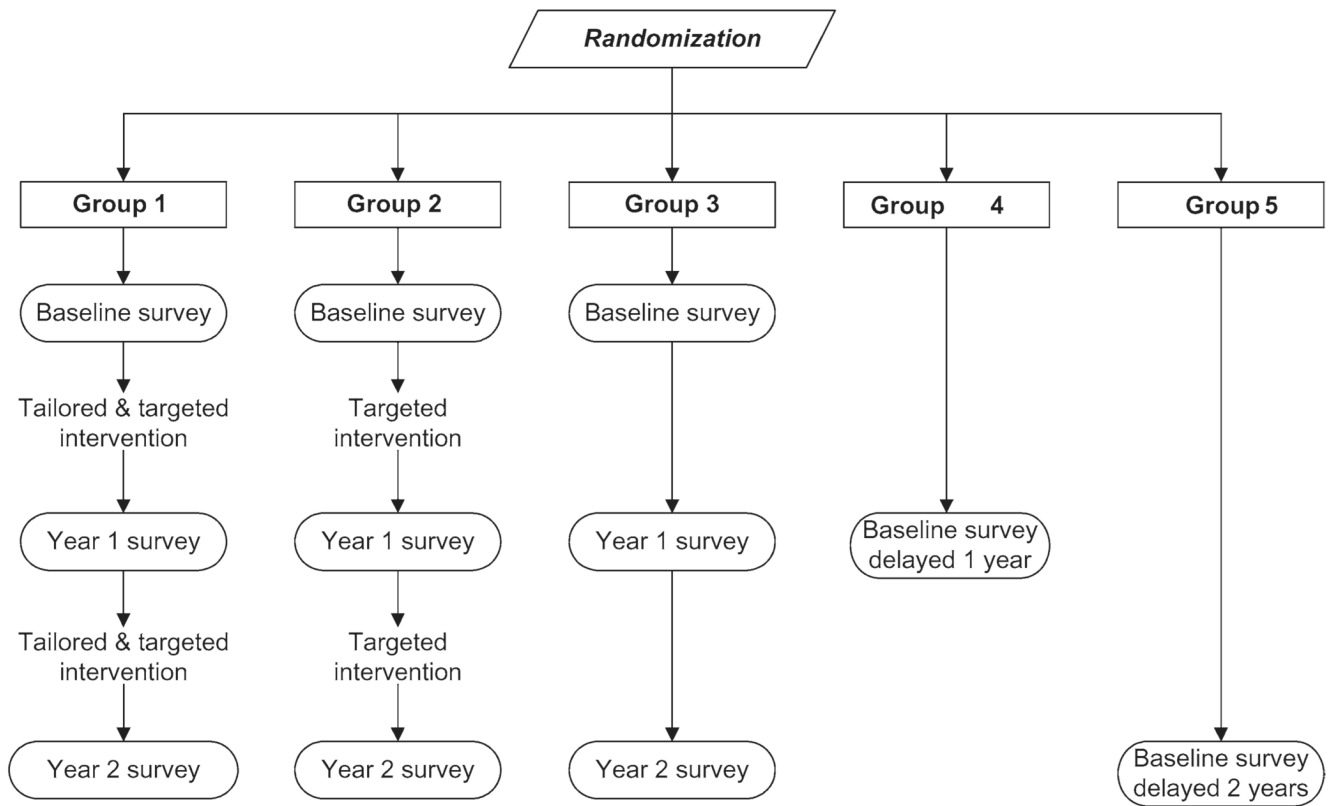


Figure 1. Research design for Project HOME (Healthy Outlook on the Mammography Experience).

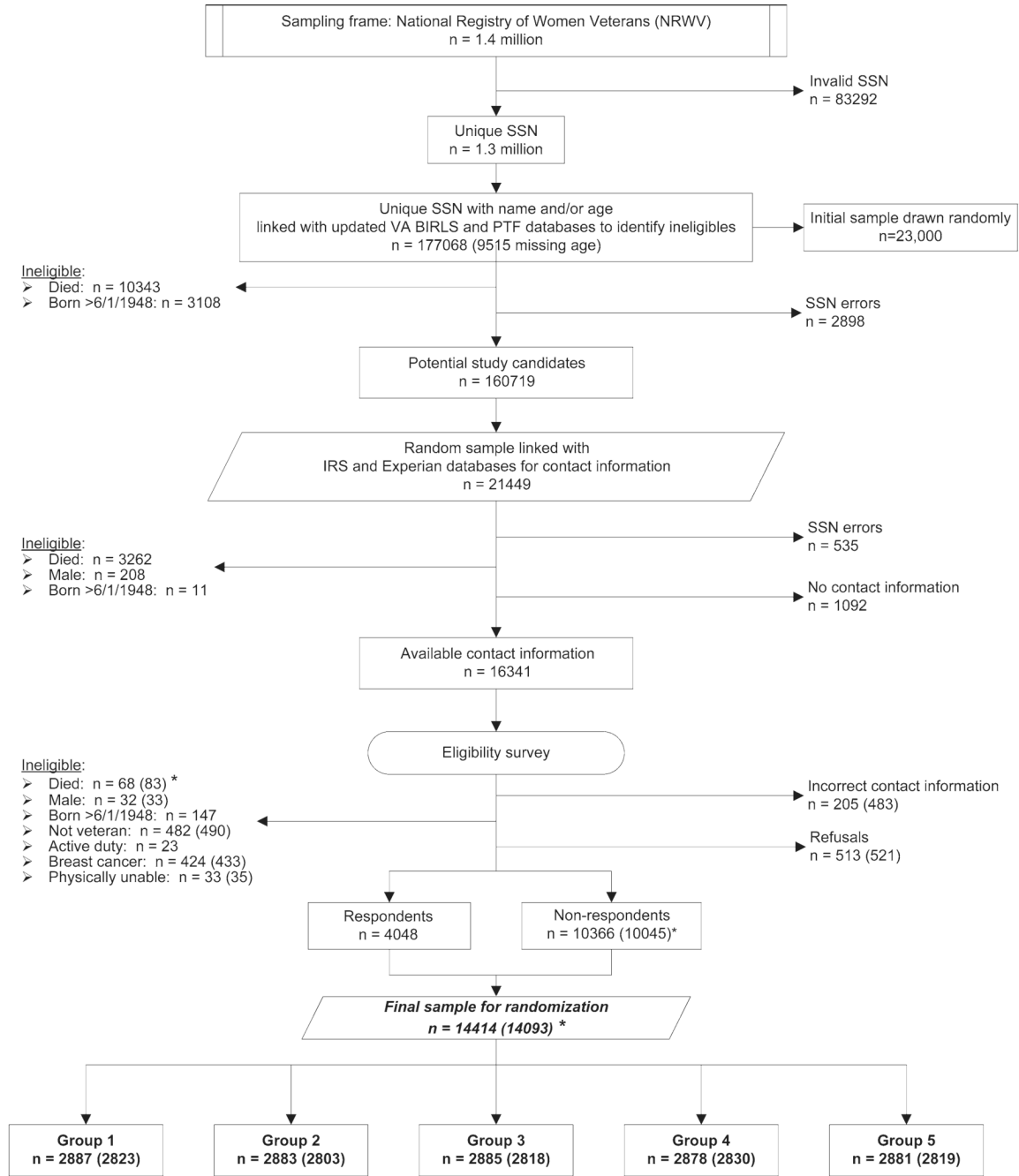


Figure 2. National Registry of Women Veterans sampling flowchart. Because differences in sample characteristics between the two sampling rounds were negligible, only the combined numbers are given. Numbers in **parentheses** reflect exclusion of subjects whose eligibility surveys were received after the randomization cutoff date and before the baseline survey was mailed. SSN = social security number; VA = Veterans Administration; BIRLS = Beneficiary Identification and Records Locator System of the VA; PTF = Patient Treatment File; IRS = Internal Revenue Service. Groups are as shown in Figure 1.

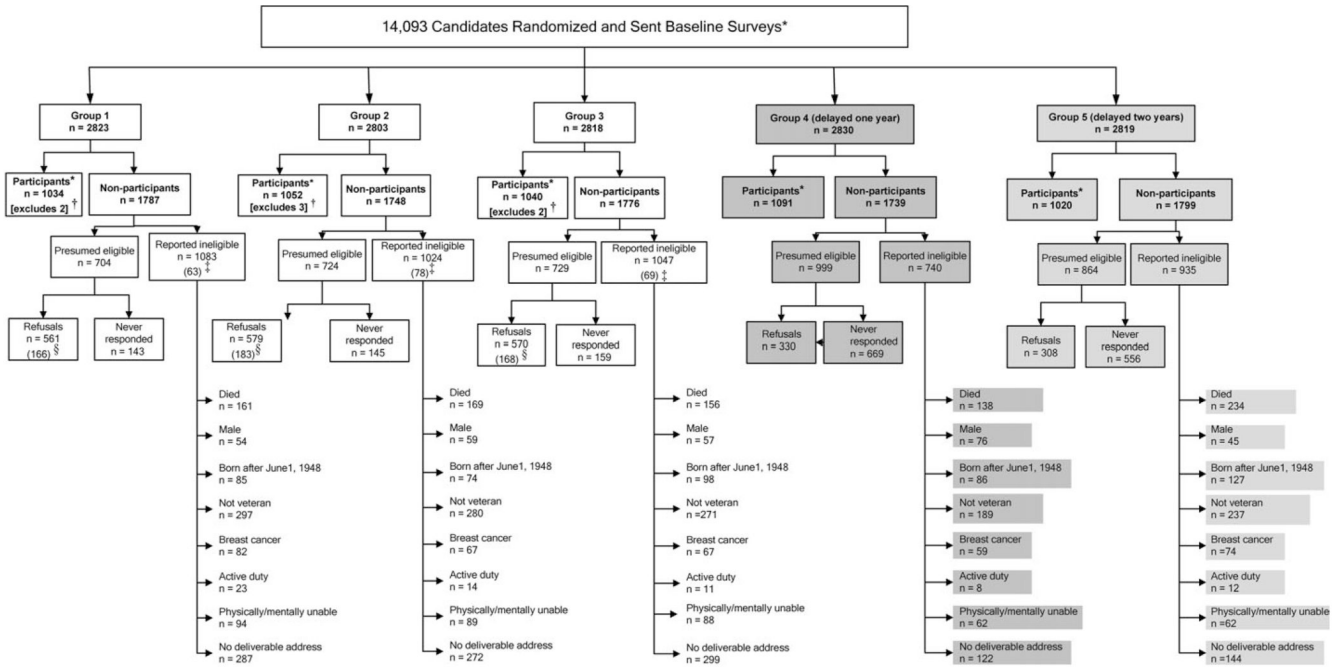


Figure 3. Flowchart of study candidates through the end of the study. For the intervention effectiveness assessment, groups 1–3 were sent parallel surveys at baseline, year 1, and year 2, whereas control group 4 (**darker shading**) was sent only the baseline survey delayed 1 year, and control group 5 (**lighter shading**) was sent only the baseline survey delayed 2 years. **Asterisk (*)** indicates that participants are candidates who provided complete mammography data on their year 1 or year 2 postintervention follow-up survey (groups 1–3) or their delayed baseline surveys (groups 4 and 5). **Dagger (†)** indicates that a total of seven participants in groups 1–3 were excluded at study’s end due to missing mammography data on the final follow-up survey. **Double dagger (‡)** indicates that reported ineligible candidates in groups 1–3 included a total of 210 candidates (**parentheses**) who became ineligible after completing a baseline survey but before the study’s end. **Section symbol (§)** indicates that withdrawals among groups 1–3 candidates included a total of 517 candidates (**parentheses**) who refused participation after completing the baseline survey. A total of 5500 study candidates in groups 1–3 were eligible for the intention-to-treat analysis of intervention effectiveness: 3126 participants and the seven candidates missing mammography data on their year 1 or year 2 follow-up survey (row 3) and 2157 presumed eligible nonparticipants and 210 reported ineligible nonparticipants who became ineligible after completing a baseline survey (row 4).

Table 1

Internal validity assessment: equivalence of study groups 1–5 at randomization (following eligibility survey)*

Independent variables	Group 1, n (%)	Group 2, n (%)	Group 3, n (%)	Group 4, n (%)	Group 5, n (%)	<i>p</i> [†]
Total	2823 (100)	2803 (100)	2818 (100)	2830 (100)	2819 (100)	
Study candidate status						.998
Eligible survey respondents	768 (27.2)	763 (27.2)	761 (27.0)	762 (26.9)	769 (27.3)	
Nonrespondents	2055 (72.8)	2040 (72.8)	2057 (73.0)	2068 (73.1)	2050 (72.7)	
Home state						.240
CA	307 (10.9)	338 (12.1)	278 (9.9)	305 (10.8)	306 (10.9)	
FL	259 (9.2)	224 (8.0)	220 (7.8)	232 (8.2)	271 (9.6)	
TX	220 (7.8)	197 (7.0)	212 (7.5)	205 (7.2)	208 (7.4)	
PA	118 (4.2)	115 (4.1)	109 (3.9)	121 (4.3)	110 (3.9)	
NY	98 (3.5)	101 (3.6)	124 (4.4)	127 (4.5)	116 (4.1)	
All others	1821 (64.4)	1828 (65.2)	1875 (66.5)	1840 (65.0)	1808 (64.1)	
NRWV source database [‡]						.370
VHA	367 (13.0)	388 (13.8)	406 (14.4)	366 (12.9)	400 (14.2)	
DoD	773 (27.4)	701 (25.0)	707 (25.1)	733 (25.9)	715 (25.3)	
BIRLS	1683 (59.6)	1714 (61.2)	1705 (60.5)	1731 (61.2)	1704 (60.5)	
Sampling round						.990
1	1520 (53.8)	1509 (53.8)	1517 (53.8)	1520 (53.7)	1513 (53.7)	
2	1303 (46.2)	1294 (46.2)	1301 (46.2)	1310 (46.3)	1306 (46.3)	
Age group, y						.610
52–64	1129 (40.0)	1093 (39.0)	1090 (38.7)	1111 (39.3)	1070 (38.0)	
≥65	1692 (60.0)	1710 (61.0)	1726 (61.3)	1719 (60.7)	1748 (62.0)	

* NRWV = National Registry of Women Veterans; VHA = Veterans Health Administration; DoD = Department of Defense; BIRLS = Beneficiary Identification and Records Locator System.

[†] From two-sided chi-square tests.[‡] VA database includes Patient Treatment File and Outpatient Clinic File; DoD database includes all separations from active duty; BIRLS database contains records of all beneficiaries, including veterans whose survivors applied for death benefits.

Table 2

Internal validity assessment: equivalence of study groups 1–5 at completion of the baseline survey*

Independent variables	Group 1, n (%)	Group 2, n (%)	Group 3, n (%)	Group 4 (delayed 1 y), n (%)	Group 5 (delayed 2 y), n (%)	P [†]
Study candidate status						.001
Nonrespondents	530 (27.1)	527 (26.3)	529 (26.8)	669 (32.0)	556 (29.5)	
Postrandomization refusals at baseline	309 (15.8)	307 (15.3)	317 (16.1)	330 (15.8)	308 (16.4)	
Eligible respondents	1116 (57.1)	1169 (58.4)	1129 (57.2)	1091 (52.2)	1020 (54.1)	<.001
Survey response mode						
Mail	960 (86.0)	968 (82.8)	981 (86.9)	997 (91.4)	906 (88.8)	
Phone	156 (14.0)	201 (17.2)	148 (13.1)	94 (8.6)	114 (11.2)	
Race/ethnicity						.342
African American	79 (7.1)	89 (7.6)	77 (6.8)	76 (7.0)	74 (7.3)	
Hispanic	42 (3.8)	25 (2.1)	29 (2.6)	42 (3.9)	29 (2.8)	
Non-Hispanic white	953 (85.4)	994 (85.0)	975 (86.4)	929 (85.2)	871 (85.4)	
Unknown	42 (3.8)	61 (5.2)	48 (4.3)	44 (4.0)	46 (4.5)	
Educational level						.687
High school or less	165 (14.8)	176 (15.1)	156 (13.8)	163 (14.9)	139 (13.6)	
Some college	471 (42.2)	512 (43.8)	518 (45.9)	467 (42.8)	449 (44.0)	
College graduate or higher	441 (39.5)	442 (37.8)	429 (38.0)	447 (41.0)	420 (41.2)	
Unknown	39 (3.5)	39 (3.3)	26 (2.3)	14 (1.3)	12 (1.2)	
Ever used VA health-care services						.138
Yes	452 (40.5)	431 (36.9)	434 (38.4)	450 (41.3)	420 (41.2)	
No	664 (59.5)	738 (63.1)	695 (61.6)	641 (58.7)	600 (58.8)	
Sampling round [‡]						.645
1	594 (53.2)	616 (52.7)	602 (53.3)	608 (55.7)	552 (54.1)	
2	522 (46.8)	553 (47.3)	527 (46.7)	483 (44.3)	468 (45.9)	
Age, y [§]						.532
52–64	728 (65.2)	789 (67.5)	726 (64.3)	720 (66.0)	681 (66.8)	
≥ 65	388 (34.8)	380 (32.5)	403 (35.7)	371 (34.0)	339 (33.2)	

* Groups 4 and 5 received the baseline survey 1 and 2 years, respectively, after groups 1–3. VA = Veterans Administration.

† From two-sided chi-square tests.

‡ Sampling round 1 occurred on September 4, 2000 (with follow-up through December 17, 2003); sampling round 2 occurred on June 1, 2001 (with follow-up through October 1, 2004).

§ Age at baseline survey mail date.

Table 3

Internal validity assessment: equivalence of study groups 1–3 at study's end*

Independent variables	Group 1, n (%)	Group 2, n (%)	Group 3, n (%)	P
Study candidate status (n = 3865)				.441
Completed baseline survey only	222 (17.4) †	259 (19.7)	232 (18.0) ‡	
Completed all three surveys	697 (55.1)	726 (55.2)	706 (55.2)	
Completed baseline + year 1 surveys	124 (9.8) §	122 (9.3) ¶	119 (9.3) ¶	
Completed baseline + year 2 surveys	73 (5.8)	62 (4.7)	72 (5.6)	
Completed year 1 and year 2 surveys	51 (4.0)	54 (4.1)	60 (4.7)	
Completed year 1 survey only	30 (2.4)	29 (2.2)	15 (1.2)	
Completed year 2 survey only	70 (5.5)	64 (4.9)	78 (6.1)	
Survey response mode # (n = 3126)				.081
Mail	812 (78.5)	834 (79.3)	833 (80.1)	
Phone	127 (12.3)	128 (12.2)	96 (9.2)	
Mail (abbreviated survey)	95 (9.2)	90 (8.6)	111 (10.7)	
Ever used VA health- care services (n = 3126)				.662
Yes	467 (45.2)	457 (43.4)	452 (43.5)	
No	567 (54.8)	595 (56.6)	588 (56.5)	
Sampling round ** (n = 3126)				.990
1	541 (52.3)	549 (52.2)	546 (52.5)	
2	493 (47.7)	503 (47.8)	494 (47.5)	
Age, y †† (n = 3126)				.677
52–64	717 (69.3)	745 (70.8)	720 (69.2)	
≥65	317 (30.7)	307 (29.2)	320 (30.8)	

* VA = Veterans Administration. P values are from two-sided chi-square tests.

† Two participants died less than 30 days after completing the baseline survey and were ineligible for the intervention effectiveness analysis.

‡ Two participants died less than 30 days after completing the baseline survey and were ineligible for the intervention effectiveness analysis.

§ Two participants had incomplete mammography data.

¶ Three participants had incomplete mammography data.

¶ Two participants had incomplete mammography data.

Three thousand one hundred and twenty-six participants had complete mammography information on their year 1 or year 2 follow-up survey.

** Sampling round 1 occurred on September 4, 2000 (with follow-up through December 17, 2003); sampling round 2 occurred on June 1, 2001 (with follow-up through October 1, 2004).

†† At baseline survey mail date.

Table 4

External validity assessment: secular trends in recent mammography screening history* at baseline survey using self-report and Veterans Health Administration (VHA) records

Independent variables	Coverage †		Compliance ‡	
	n (%)	<i>p</i> §	n (%)	<i>p</i> §
Baseline survey respondents (self-report) (n = 5525)				
Group 1	947 (84.9)	.641	647 (58.0)	.538
Group 2	1005 (86.0)	.219	698 (59.7)	.825
Group 3 (referent)	950 (84.2)	Referent	669 (59.3)	Referent
Group 4 (baseline delayed 1 y)¶	931 (85.3)	.436	618 (56.7)	.213
Group 5 (baseline delayed 2 y)¶	839 (82.3)	.241	604 (59.2)	.985
Subgroup of those who ever used VA health-care services, including baseline survey respondents, nonrespondents, and refusals (VA records) ¶ (n = 3676)				
Group 1	169 (21.8)	.478	89 (11.5)	.689
Group 2	159 (21.9)	.521	87 (12.0)	.486
Group 3 (referent)	166 (23.3)	Referent	77 (10.8)	Referent
Group 4 (baseline delayed 1 y)¶	176 (22.7)	.781	82 (10.6)	.884
Group 5 (baseline delayed 2 y)¶	118 (17.2)	.004	61 (8.9)	.228

* Recent mammography screening history was defined as having had one or more mammograms within the 30 months preceding the date of each group's baseline survey. For self-report, baseline survey respondents reported the dates of their two most recent mammograms. For VHA records, mammogram procedure codes and dates between the years 1996 and 2005 were documented in Veterans Administration (VA) administrative databases, either the Patient Treatment File or the Outpatient Clinic File, for the subgroup of those who had ever used VA health-care services.

† Coverage = at least one prior mammogram.

‡ Compliance = at least two prior mammograms 6–15 months apart.

§ From two-sided chi-square tests.

¶ Groups 4 and 5 received the baseline survey 1 and 2 years, respectively, after groups 1–3.

¶ Excludes decedents, males, and those born after June 1, 1948, as ascertained by linkages with available electronic databases.

Table 5

External validity assessment: adjusted ORs (with 95% CIs) comparing refusing and never responding with participating (n = 3126) in groups 1–3 at study's end*

Independent variables	Refusal (n = 1710)	Never responded (n = 447)
Group		
1	0.99 (0.86 to 1.15)	0.90 (0.71 to 1.15)
2	1.01 (0.88 to 1.17)	0.90 (0.71 to 1.14)
3	Referent	Referent
NRWV source database [†]		
VHA	1.31 (1.11 to 1.54)	1.50 (1.14 to 1.98)
DoD	1.21 (1.02 to 1.43)	1.22 (0.92 to 1.60)
BIRLS	Referent	Referent
Sampling round [‡]		
1	0.98 (0.87 to 1.11)	0.62 (0.51 to 0.76)
2	Referent	Referent
Age group, y [§]		
52–64	0.59 (0.52 to 0.68)	1.32 (1.04 to 1.68)
≥ 65	Referent	Referent

* Group-specific numbers are as given in Figure 3. Odds ratios (ORs) and 95% confidence intervals (CIs) were adjusted for all other variables (ie, group, National Registry of Women Veterans [NRWV] source database, sampling round, or age group) using multiple logistic regression. VHA = Veterans Health Administration; DoD = Department of Defense; BIRLS = Beneficiary Identification and Records Locator System.

[†] VA database includes Patient Treatment File and Outpatient Clinic File; DoD database includes all separations from active duty; BIRLS database contains records of all beneficiaries, including veterans whose survivors applied for death benefits.

[‡] Sampling round 1 occurred on September 4, 2000 (with follow-up through December 17, 2003); sampling round 2 occurred on June 1, 2001 (with follow-up through October 1, 2004).

[§] At baseline survey mail date.