

## SUPPLEMENT

The time-to-event model is a general form of a Cox proportional hazards (CPH) model expressed as,  $h(t) = h_0(t)e^{b_1X_1+b_2X_2+\dots+b_nX_n}$ . Where  $h$  is the expected hazard at a time  $t$  and  $h_0(t)$  is the baseline hazard when the predictors ( $X_1, X_2, \dots, X_n$ ) equal zero [1]. The simulations had four events: a) National Breast and Cervical Cancer Early Detection Program (Program) entry, b) receiving a mammogram, c) receiving a breast cancer diagnosis, and d) death. Data from the Program's administrative database called Minimum Data Elements (MDEs), the National Health Interview Survey (NHIS) [2], the United States Cancer Statistics (USCS) [3], and data from the literature were used as input parameters in the model [4,5]. Figure S1 presents an expanded model schematic of the methods.

### *Program growth and demographics*

We used MDEs Version 7.0 for Program growth and participant demographics to develop a model that simulates patients entering and exiting the Program. For each woman who participated in the Program from 2010 to 2019, we simulated the events of a participant from age 40 years to death, or the maximum age of 100 years, with a 6-month time step (cycle length). From MDEs, we selected participants born only between 1946 and 1976—because if a participant were born in 1946, they would have been age 64 years at the start of our time horizon in 2010. If a participant were born in 1979, they would have been age 40 years at the end of our time horizon in 2019. After selecting participants who had received a mammogram through the Program from 2010 to 2019, we had 3,288,096 unique individuals within the dataset.

We stratified MDE by race/ethnicity, but we excluded from the analysis participants with at least one unknown racial record (60,682 participants, 1.90%) and those who reported more than one race/ethnicity (162 participants, 0.005%). The remaining participants were grouped according to these categories, non-Hispanic White (NHW), non-Hispanic Black (NHB), non-Hispanic American Indian/Alaskan Native (AI/AN), and non-Hispanic Asian or Pacific Islander (NH API) persons. For the years 2010–2019, there were 997,478 (16.09%) NHB participants, 2,646,410 (42.68%) NHW participants, 590,084 (9.52%) NH AIAN/API and 1,966,490 (31.72%) Hispanic participants. Of the NH AIAN/API participants, 51.08% were identified as Asian, 41.28% were identified as NH AIAN, 4.85% were identified as Hawaiian/Pacific Islander, and 2.79% were identified as other API.

Based on the data, we observed that the overall annual number of screenings provided by the Program to new and existing participants of all races/ethnicities had changed over time (Figure S2) with more screening from 2010 through 2012 and slowly declining from 2014 to 2019. It was assumed that the declining trend observed was

due to the passage of the Affordable Care Act (ACA) in 2010 and the subsequent expansion of Medicaid in 2014 [6,7]. Figure S3 shows the pattern of breast cancer screening decline by race/ethnicity for new (panel A) and existing (Panel B) participants. For example, in 2010, new participants were predominantly NHW, but in 2014-2015, Hispanic women became the largest population of new Program participants.

To model Program growth, we classified participants as existing or new, depending on when they received their first screening from the Program. At the beginning of our simulations, we defined a participant as “existing” if the participant received a screening from the Program before 2010. Participants were defined as “new” when they received their first screening from the Program. A new participant becomes an existing participant in the next cycle. A generalized additive model was used to model the number of new and existing participants, and the R code used was: `gam(y~s(age,k=N_age,bs="cr," pc=ref_age), family=" Poisson," knots=list(age=ages), gamma=gamma, control=list(trace=trace)).`

### ***Screening initiation and behaviors***

The weighted percentage of women who received a mammogram by age was determined using 2010, 2013, 2015, 2018, and 2019 NHIS responses to the question, “Have you ever received a mammogram.” The respondents were categorized into insured and uninsured based on whether the respondent had private or public insurance or neither, and they were also grouped by 5-year age-group (40–44, 45–49, 50–54, 55–59, 60–64, 65–69, and 70+ years). We excluded missing values and respondents enrolled in Medicaid. NHIS data were analyzed from respondents <250% of the federal poverty level (FPL). The insured respondents were the proxy for Program participants, and those uninsured were the counterfactual. The weighted percent (and 95% confidence interval) of women who had ever had a mammogram for each response year were pooled for 2010, 2013, and 2015. Data from 2018 and 2019 were not pooled because of a different survey design in those years.

The percentage of insured respondents who had ever received a mammogram at the age group 60–64 years was 94.61%. Because we assumed that all participants in the Program (insured NHIS respondents) had at least one screening by age <65, we normalized the Program participants to all have ever received a mammogram at age 60–64 years and divided all the uninsured percentages by 94.61%. Then, we divided the age groups into 6-month increments. We assumed that the percentage of women who had ever received a mammogram increased equally every 6 months for the age group until the percentage of women who had ever received a mammogram was reached at the maximum age of the age group. For example, if the percentage of women in the age group 40–44 years who

had ever had a mammogram was 50%, then we would have designated the percentage of women aged 40 years who had ever had a mammogram to be 5%, aged 40.5 years to be 10%, and so on until the percentage of women aged 44.5 years who had ever had a mammogram was designated to be 50%. After 65 years of age, the percentage of women who had ever received a mammogram did not change.

We assumed that a participant might start screening at Program entry or following the percentage of women who had ever had a mammogram by 6-month increments as defined by data from NHIS respondents. If a woman entering the Program received a mammogram before, based on the individual's age and the NHIS respondent data described above, then the age of the first mammogram was defined as when they received their mammogram outside of Program participation. If a woman had not received a mammogram before, we assumed their first mammogram was upon Program entry.

After Program entry, we analyzed the screening behaviors among participants who had at least one mammogram before 2016 using the MDE data. Therefore, there was at least a 5-year window to observe the participants' screening behaviors before making predictions about their future behaviors. We estimated the proportion of participants who only had one mammogram through the Program by age at Program entry (age 40–49, 50–59, 60–64 years) and by race/ethnicity. The proportion of one-time screeners by race/ethnicity is presented in Figure S4. NHB and NHW participants had the highest proportions of one-time screeners.

For participants who received more than one screening through the Program, the mean duration between mammograms was calculated and used to classify participants into the following screening behaviors: annual, biennial, and irregular. The mean duration between mammograms was also used to calculate the 6-month screening rates and probabilities. The screening probabilities were used in the simulation model. Individual screening behaviors changed over time as a woman moved into the next age group through the simulation model. For example, if a woman was initially in the 50–54-year age group and that woman turned 55 years of age, the woman then followed the screening behaviors defined by the 55–59-year age group.

We estimated the reduction in breast cancer incidence for Program participants compared to women who were uninsured with annual and biennial screening behaviors from supplementary data in Table 1 [4]. Table 6 provides the annual, biennial, and irregular screeners percentage among uninsured women, stratified by age. A reduction factor was calculated by dividing the screening behavior by age in the Program and the screening behavior by age for uninsured women. Then, the screening behavior by age group and race/ethnicity for Program participants

was divided by the reduction factor to determine the proportion of uninsured women in each screening behavior category. The proportions of Program participants and uninsured women were categorized by screening behavior presented by race/ethnicity in Figure S5.

### ***Breast cancer diagnosis***

For each race/ethnicity, we first calculated the average incidence of breast cancer per age group per cycle. For women aged 85 years or older, we assumed that the incidence rate was the same for up to 100 years. Then, the incidence rate was converted to a probability for each age group per cycle. We followed a lifetable approach to calculate the time to breast cancer diagnosis, or survival curve, for each race/ethnicity for women in the insured category. The lifetable calculations served as the baseline survival functions by race/ethnicity. Then, the survival curves for Program participants were calculated by using the coefficients from the Cox proportional hazard regression. The R code example for this is: `coxph(Surv(age_center, event_dx) ~ mde_ins, method = "breslow", data = survdf[race == "White"], x = T)`.

We used the USCS data for women aged 40 years or older to estimate the hazard ratio and the proportion of ductal carcinoma in situ, local, regional, and distant breast cancer diagnoses for each race/ethnicity. We used the national age-specific female breast cancer incidence rate by race/ethnicity from January 1, 2010, until December 31, 2017, as the incidence for women in the insured category. For each race/ethnicity, a Cox proportional hazards model was used to calculate the probability of receiving a breast cancer diagnosis at each cycle. Instead of the probability of survival being the outcome of the Cox proportional hazards model, the outcome was the probability of not receiving a breast cancer diagnosis. The baseline hazard was designated to be the breast cancer diagnosis incidence for women with insurance. The R code for the Cox proportional hazards model is `cox_RACE <- coxph(Surv(age, breastCancerDiagnosis) ~ ProgramParticipant, method = "breslow," data = USCS, x = T)`.

Then, hazard functions were calculated for Program participants and uninsured women. Using the Cox proportional hazard regression models, we also estimated the hazard ratio of breast cancer diagnosis among uninsured women relative to insured women and for each race/ethnicity. The probability of not receiving a breast cancer diagnosis from the hazard functions by race/ethnicity and by Program versus no Program was used to predict whether a woman received a breast cancer diagnosis in the simulation model.

The stage of breast cancer diagnosis by race/ethnicity and by Program versus no Program was also calculated from USCS data and used in the simulation model. A summary of the proportions of breast cancer stage by age range is shown in Figure S6.

### ***Survival***

The R code to estimate the hazard ratio of breast cancer survival relative to Program status is `cox_RACE <- coxph(Surv(time, status) ~ mde_ins + stage + mde_ins * stage + age_dx * stage, method = "Breslow," data = USCS, x = T)`. To calculate the survival of women who did not have a breast cancer diagnosis, we followed the methods described by Trentham-Dietz et al. and Gangnon et al., which estimated the hazard ratio for breast cancer survival up to age 100 years [5,8].

### ***Sensitivity analysis***

For the proportions of annual, biennial, and irregular screeners to add up to one, when a proportion of a screening behavior was changed for the sensitivity analysis, the other proportions were adjusted. For example, increasing the proportion of annual screeners resulted in a reduction of biennial and irregular screeners. Decreasing the proportion of annual screeners resulted in a proportional increase in biennial and irregular screeners. Increasing the proportion of biennial screeners was assumed to be due to irregular screeners screening more frequently and not because annual screeners screened less. So, the proportion of annual screeners did not change, and the proportion of irregular screeners decreased by 10%. Reducing the proportion of biennial screeners by 10% caused an increase in the proportion of irregular screeners by 10%. Increasing the proportion of irregular screeners was assumed to result from an overall decrease in the number of screenings received by Program participants. Hence, the proportion of annual screeners decreased, and the proportion of biennial screeners increased due to annual screeners screening less frequently. Decreasing the proportion of irregular screeners increased the proportion of annual screeners and decreased the proportion of biennial screeners due to all participants screening more frequently. The exact proportions used in each sensitivity analysis are presented in Table S1.

### ***Simulation***

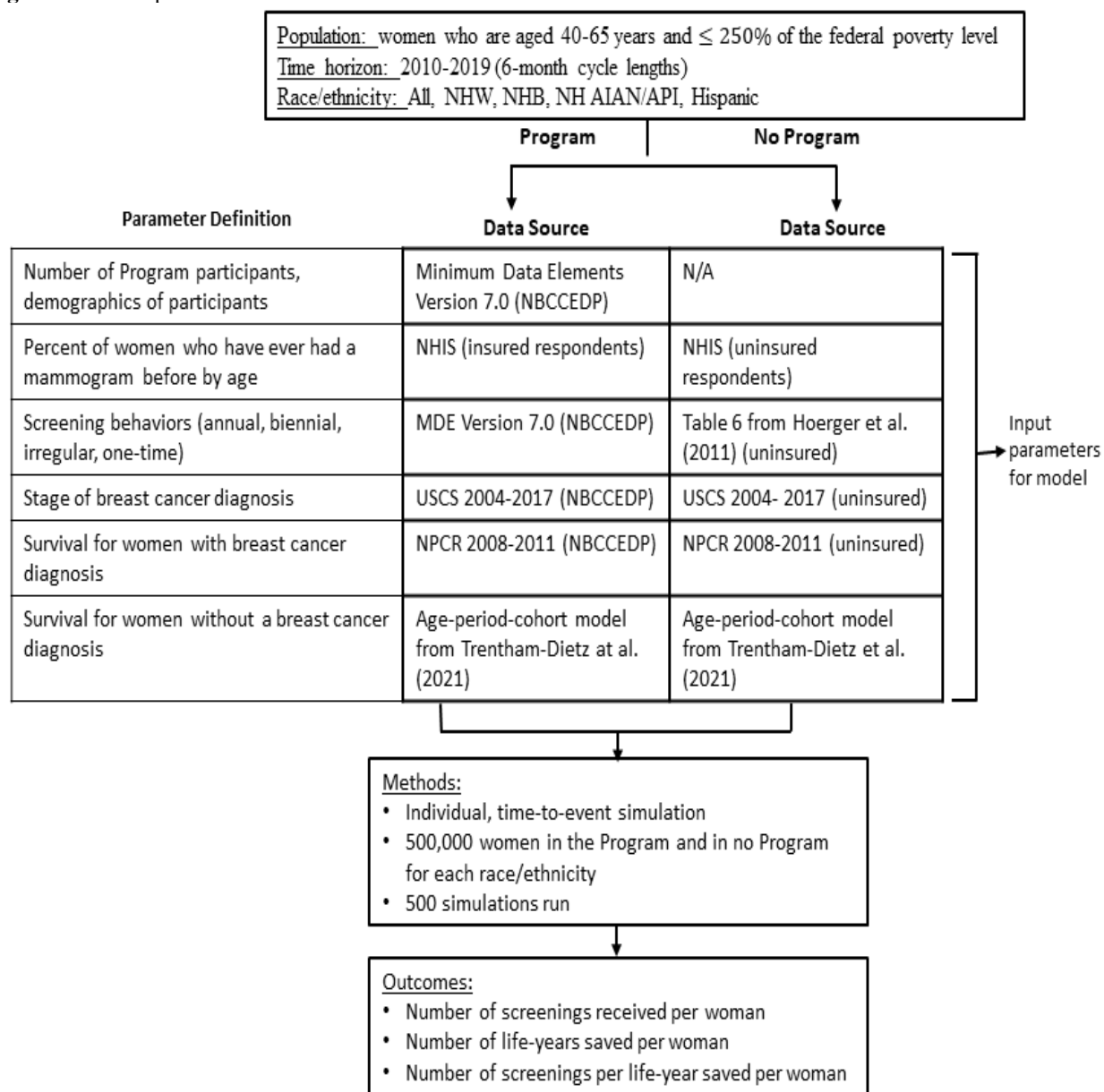
The simulations in Python (Python Software Foundation, Beaverton, Oregon) used a random seed to sample by assigned random percentile. The screening behavior, stage of breast cancer diagnosis, and age at death were assigned to women within the simulation based on random percentile sampling. After 500 simulations with 500,000 women in each simulation, the results were averaged and scaled to be reported for 1 million women.

## REFERENCES to SUPPLEMENT

1. Deo SV, Deo V, Sundaram V. (2021) Survival analysis-part 2: Cox proportional hazards model. *Indian J Thorac Cardiovasc Surg.* 37: 229-33.
2. Analysis of National Health Interview Surveys, 2010-2019: NHIS Data, Questionnaires and Related Documentation. <https://www.cdc.gov/nchs/nhis/data-questionnaires-documentation.htm>. Accessed 8 August 2024
3. U.S. Cancer Statistics Working Group (2022) U.S. Cancer Statistics Data Visualizations Tool, based on 2021 submission data (1999-2019): U.S. Department of Health and Human Services, Centers for Disease Control and Prevention and National Cancer Institute. <https://www.cdc.gov/cancer/dataviz>. Accessed 5 August 2024.
4. Hoerger TJ, Ekwueme DU, Miller JW et al (2011) Estimated effects of the National Breast and Cervical Cancer Early Detection Program on breast cancer mortality. *Am J Prev Med.* 40: 397-404.
5. Trentham-Dietz A, Chapman CH, Bird J, Gangnon RE (2021) Recent changes in the patterns of breast cancer as a proportion of all deaths according to race and ethnicity. *Epidemiology.* 32: 904-13.
6. Tangka F, Kenny K, Miller J, Howard DH (2020) The eligibility and reach of the national breast and cervical cancer early detection program after implementation of the Affordable Care Act. *Cancer Causes & Control.* 31: 473-89.
7. Wong FL, Miller JW (2019) Centers for disease control and prevention's national breast and cervical cancer early detection program: increasing access to screening. *J. women's health.* 28: 427-31.
8. Gangnon RE, Stout, NK, Alagoz O, Hampton JM, Sprague BL, Trentham-Dietz A. (2018). Contribution of breast cancer to overall mortality for US women. *Medical Decision Making*, 38(1\_suppl), 24S-31S.

## SUPPLEMENT – TABLES AND FIGURES

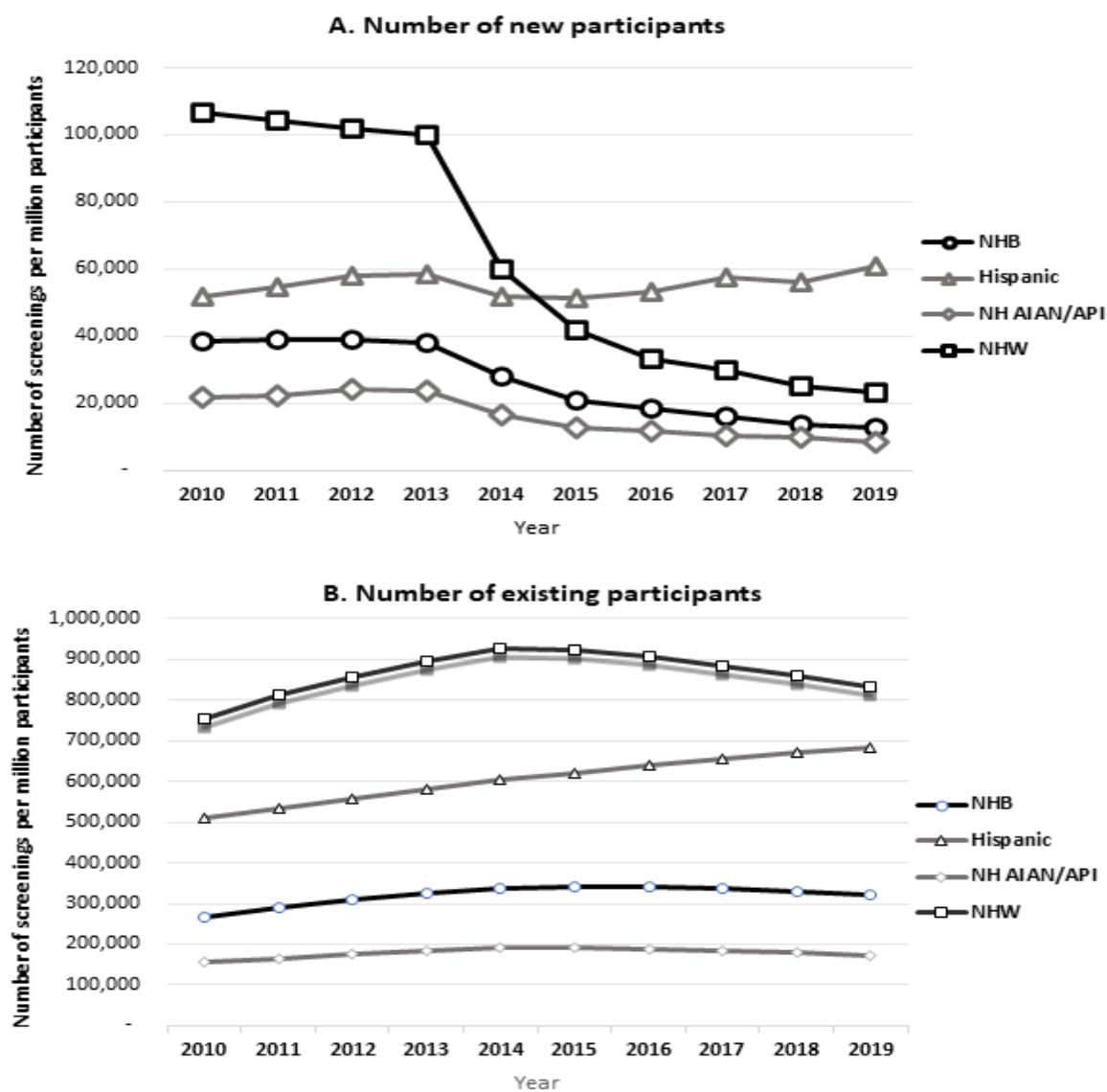
**Figure S1.** An expanded model schematic.\*



Abbreviations: MDEs, Minimum Data Elements; NBCCEDP, National Breast and Cervical Cancer Early Detection Program; NHB, non-Hispanic Black; NHW, non-Hispanic White; American Indian or Alaska Native, Asian, Native Hawaiian, or other Pacific Islander; NHIS, National Health Interview Survey; NPCR National Program of Cancer Registries; USCS, United States Cancer Statistics

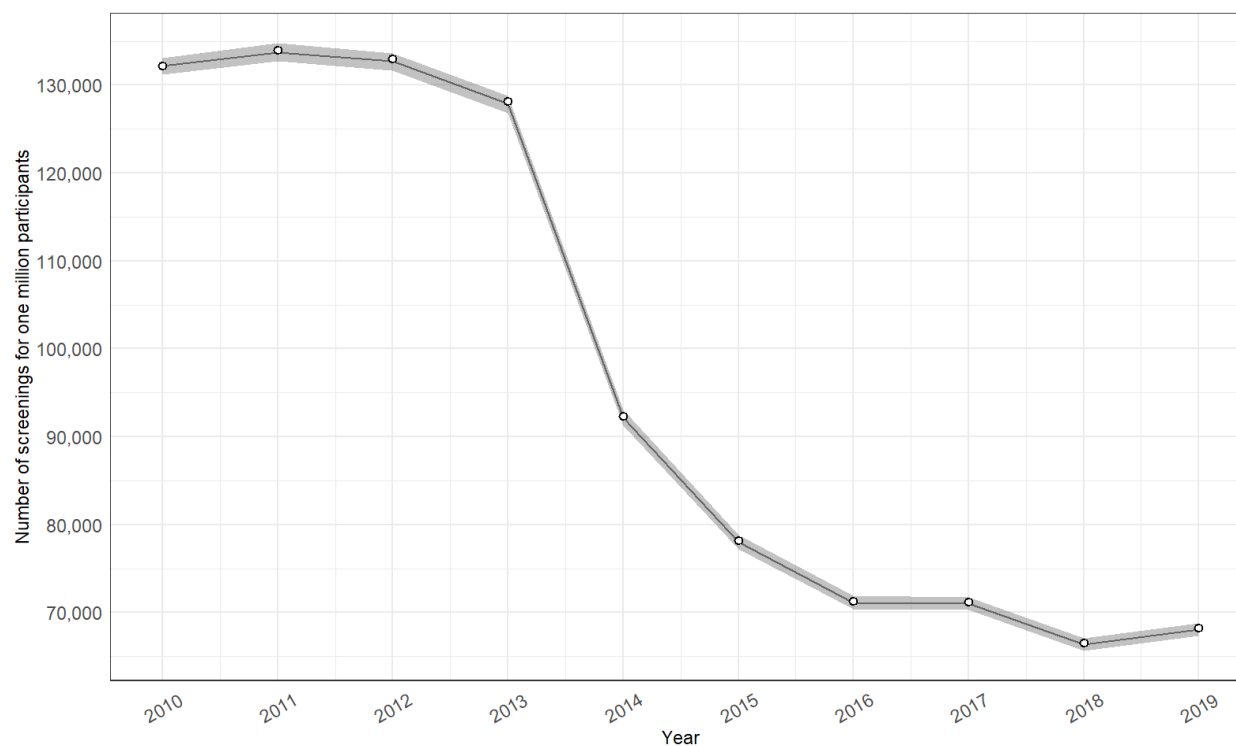
\*The top box provides context for the data collected, the data source columns are the sources of data for Program (left) and no Program (right) for the parameter defined on the far left (Parameter Definition).

**Figure S2.** The Program’s annual number of new and existing breast cancer screening per million participants by race/ethnicity, 2010—2019



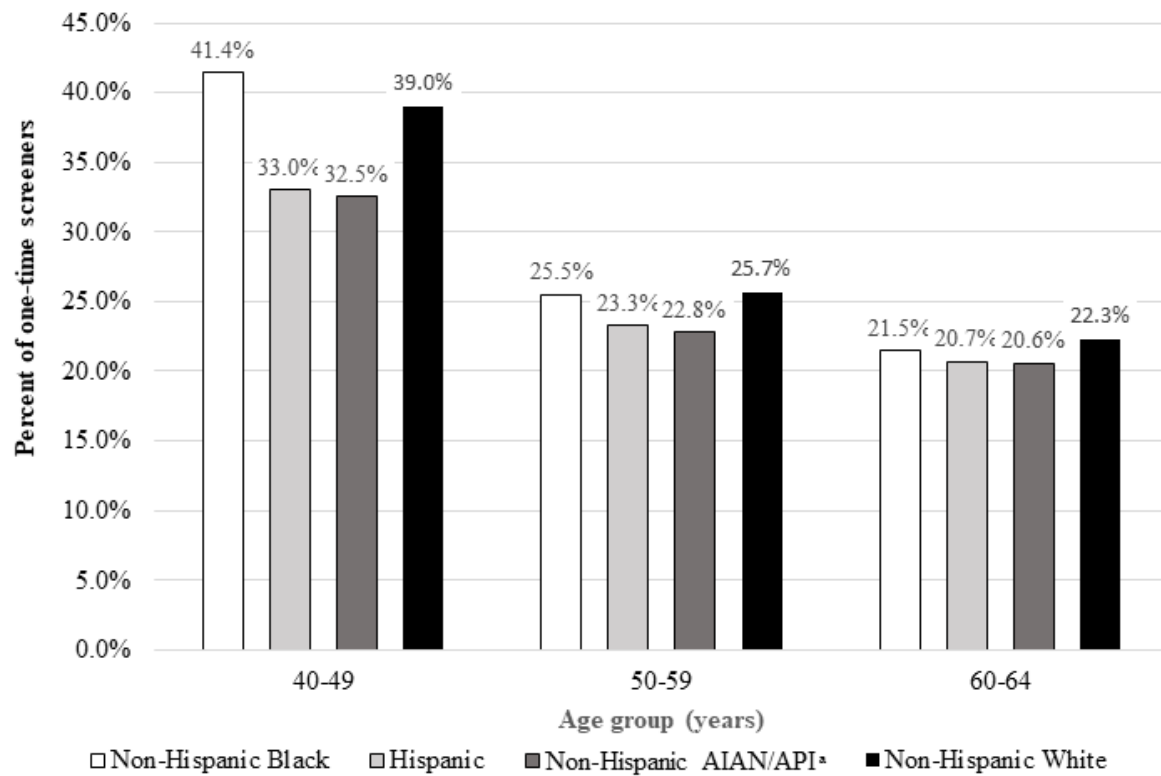
Abbreviations: NHB, non-Hispanic Black; NH-AIAN/API, those who identified as American Indian or Alaska Natives, and Asian, Native Hawaiian, or other Pacific Islander; NHW, non-Hispanic White.

**Figure S3.** The annual number of screenings provided by the NBCCEDP among 1 million participants, 2010—2019<sup>a</sup>



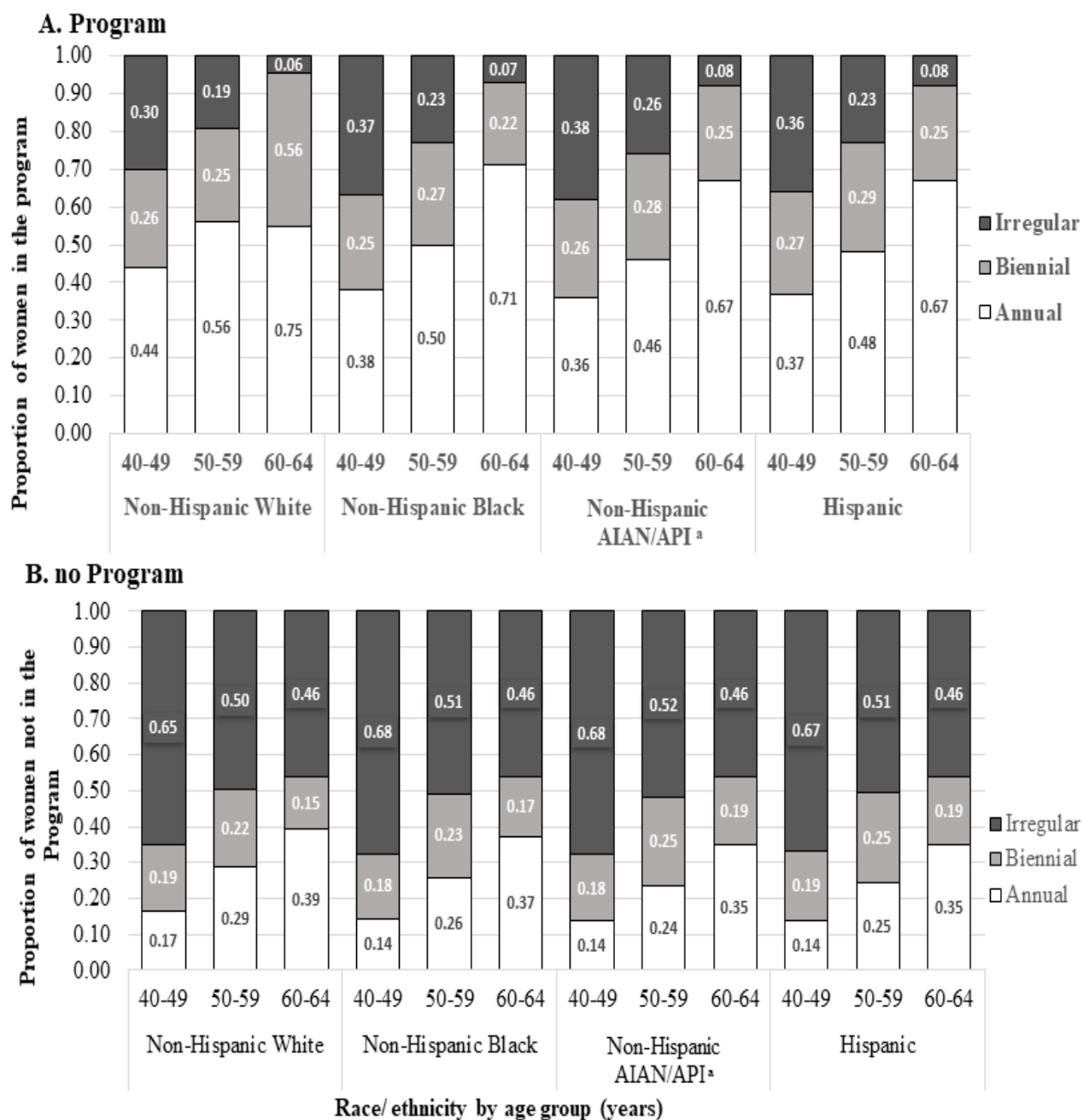
<sup>a</sup> The data points from the Program's MDEs database are represented by the dots, the simulation results are represented with the line, and the shaded area around the line is the 95% confidence interval for the simulation results

**Figure S4.** The percentage of Program participants who were one-time screeners varied by race/ethnicity, 2010—2019



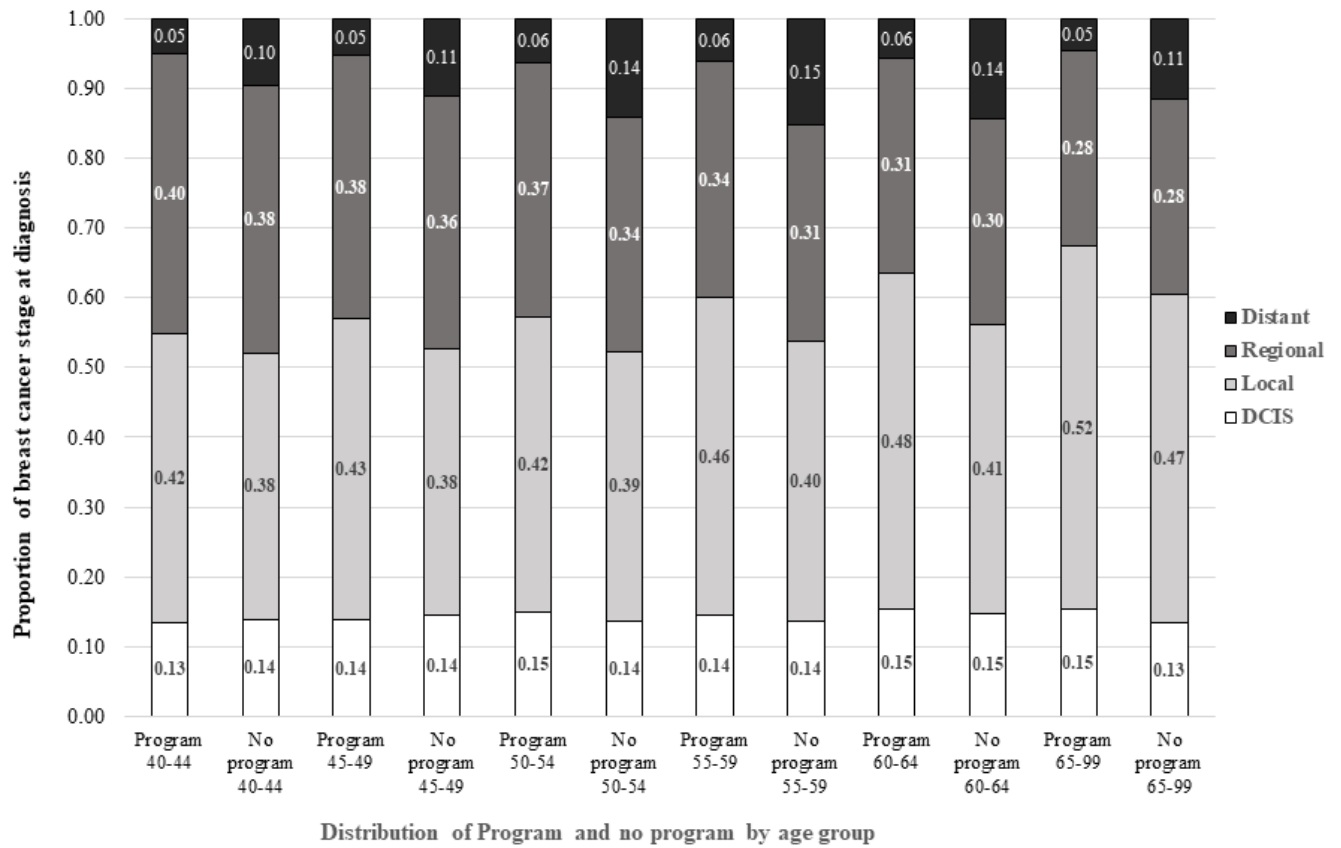
<sup>a</sup> Non-Hispanic AIAN/API race/ethnicity includes those who identify as American Indian or Alaska Native, Asian, Native Hawaiian, or other Pacific Islander. These racial/ethnic groups were combined because of their small sample size. However, NBCCEDP supports each tribe and tribal organization's program.

**Figure S5.** The proportion of women in the Program and no Program with different levels of screening behavior by race/ethnicity, 2010–2019



<sup>a</sup> Non-Hispanic AIAN/API race/ethnicity includes those who identify as American Indian or Alaska Native, Asian, Native Hawaiian, or other Pacific Islander. These racial/ethnic groups were combined because of their small sample size. However, NBCCEDP supports each tribe and tribal organization's program.

**Figure S6.** Distribution of breast cancer diagnosis stages among Program participants and women who were uninsured by age range and by all racial/ethnic groups, 2010—2019



Abbreviation: DCIS, ductal carcinoma in situ.

**Table S1.** Proportions of Program participants by screening behavior for the sensitivity analysis, 2010—2019

| Proportions of Program participants                   | Annual | Biennial | Irregular |
|-------------------------------------------------------|--------|----------|-----------|
| Base case proportions                                 | 0.496  | 0.260    | 0.244     |
| Increase the proportion of annual screeners by 10%    | 0.596  | 0.214    | 0.190     |
| Decrease the proportion of annual screeners by 10%    | 0.396  | 0.301    | 0.303     |
| Increase the proportion of biennial screeners by 10%  | 0.496  | 0.360    | 0.144     |
| Decrease the proportion of biennial screeners by 10%  | 0.496  | 0.160    | 0.344     |
| Increase the proportion of irregular screeners by 10% | 0.333  | 0.323    | 0.344     |
| Decrease the proportion of irregular screeners by 10% | 0.680  | 0.177    | 0.144     |

These proportions were obtained from Heorger et al [4].

In this table, the term “screeners” refers to those women who received breast cancer screening either annually, biennially, or irregularly.