

UC Berkeley

UC Berkeley Previously Published Works

Title

Hypothetical Interventions on Diet Quality and Lifestyle Factors to Improve Breast Cancer Survival: The Pathways Study.

Permalink

<https://escholarship.org/uc/item/9ht8r35x>

Journal

Cancer Epidemiology, Biomarkers & Prevention, 32(12)

Authors

Ergas, Isaac

Bradshaw, Patrick

Cespedes Feliciano, Elizabeth

et al.

Publication Date

2023-12-01

DOI

10.1158/1055-9965.EPI-22-1216

Peer reviewed



Published in final edited form as:

Cancer Epidemiol Biomarkers Prev. 2023 December 01; 32(12): 1716–1725.

doi:10.1158/1055-9965.EPI-22-1216.

Hypothetical interventions on diet quality and lifestyle factors to improve breast cancer survival: the Pathways Study

Isaac J. Ergas^{1,2}, Patrick T. Bradshaw², Elizabeth M. Cespedes Feliciano¹, Janise M. Roh¹, Marilyn L. Kwan¹, Barbara Laraia^{2,3}, Kristine A. Madsen³, Song Yao⁴, Catherine Thomsen⁵, Lawrence H. Kushi¹

¹Division of Research, Kaiser Permanente Northern California, 2000 Broadway, Oakland, CA 94612

²Division of Epidemiology, School of Public Health, University of California, Berkeley, CA 94720

³Division of Community Health Sciences, School of Public Health, University of California, Berkeley, CA 94720

⁴Department of Cancer Prevention and Control, Roswell Park Comprehensive Cancer Center, Buffalo, NY 14263

⁵Zero Breast Cancer, San Rafael, CA 94903

Abstract

Background: The number of breast cancer survivors is increasing, yet evidence to inform dietary and lifestyle guidelines is limited.

Methods: This analysis included 3,658 participants from the Pathways Study, a prospective cohort of women diagnosed with invasive breast cancer. A healthy plant-based dietary index score (hPDI), an American Cancer Society nutrition guidelines score (ACS), a 2015 Healthy Eating Index score (HEI), hours per week of moderate to vigorous physical activity (PA) and lifetime cumulative pack-years of cigarette smoking (SM) were each measured at diagnosis, 6, 24 and 72 months. Using g-computation, 5- and 10-year risk ratios (RRs), risk differences (RDs), and 95% confidence intervals (CIs) for all-cause mortality under hypothetical interventions on diet quality, physical activity, and smoking, compared to the natural course (no intervention) were calculated.

Results: Hypothetical moderate to extreme interventions on hPDI, ACS and HEI, each in combination with PA and SM, showed 11 to 56%, 9 to 38%, and 9 to 49% decreases in 5-year risks of all-cause mortality compared to no intervention, respectively [(hPDI: $RR_{\text{moderate}}=0.89$, 95% CI 0.82-0.94; $RR_{\text{extreme}}=0.44$, 95% CI 0.26-0.67), (ACS: $RR_{\text{moderate}}=0.91$, 95% CI 0.85-0.96; $RR_{\text{extreme}}=0.62$, 95% CI 0.43-0.82), (HEI: $RR_{\text{moderate}}=0.91$, 95% CI 0.84-0.95; $RR_{\text{extreme}}=0.51$, 95% CI 0.33-0.72)]. While 10-year relative risks were slightly attenuated, absolute risk reductions were more pronounced.

Corresponding author Isaac J. Ergas, Division of Research, Kaiser Permanente Northern California, 2000 Broadway, Oakland, CA 94612, isaac.j.ergas@kp.org, Tel: 510-813-3366.

Author disclosures

The authors declare no potential conflicts of interest.

Conclusions: Interventions to improve diet quality, increase physical activity, or reduce smoking at the time of diagnosis may improve survival among breast cancer survivors.

Impact: We estimate that over 10% of deaths could be delayed by even moderate adoption of these behaviors.

Keywords

Breast cancer; Survivors; Mortality; Diet; Physical activity; Smoking; G-formula

INTRODUCTION

Excluding skin-cancers, breast cancer is the most commonly-diagnosed cancer among woman globally.¹ However, female breast cancer survivors are now living longer than ever, and there are currently over 3.8 million survivors in the United States (US).² These women are highly motivated to make lifestyle changes, and desire more evidence-based information from health professionals.³⁻⁵ Nevertheless, there is limited evidence to inform recommendations for this population, therefore they are typically offered guidance based on prevention, not survival.^{6, 7}

Associations of diet quality, physical activity, and smoking with all-cause mortality after breast cancer have each been independently shown,⁷⁻⁹ yet their effects in combination have been less studied.^{6, 7} Dietary patterns (a composite measure of the quantities, portions, and frequencies of one's dietary intakes)¹⁰ concordant with healthy eating have been associated with reduced mortality among breast cancer survivors,¹¹⁻¹⁴ as has physical activity both before and after diagnosis.¹⁵⁻¹⁷ While few studies have examined smoking at or around the time of diagnosis, those that have, showed smoking to negatively impact long-term survival.^{8, 18}

Though lifestyle recommendations for cancer patients would be best informed by evidence from a randomized controlled trial (RCT), there are few such studies.^{19, 20} Alternatively, causal inference methods such as parametric g-computation can be applied to observational data to estimate and compare risks under hypothetical interventions on risk factors either individually or in combination.²¹⁻²³ While this approach does not do away with confounding from unknown sources (as might randomization in a properly designed RCT), when assumptions of no unmeasured confounding, no measurement error, no model misspecification, causal consistency, and positivity are met, it does facilitate appropriate adjustment for time-varying confounders affected by prior exposures that might otherwise lead to bias when using standard statistical methods.²¹⁻²³

Thus, we applied parametric g-computation to data from the Pathways Study, a prospective cohort of women with breast cancer, to estimate the risk of all-cause mortality under hypothetical interventions on dietary quality, physical activity, and smoking at the time of breast cancer diagnosis. To our knowledge, this study is the first to evaluate the combined effect of diet and lifestyle factors on survival among women with breast cancer using a rigorous causal inference framework.

MATERIALS AND METHODS

Study Cohort

A total of 4,505 female breast cancer survivors diagnosed with invasive breast cancer between the years 2005 and 2013 at Kaiser Permanente Northern California (KPNC), were enrolled in the Pathways Study. Study participants were recruited from a cohort of 11,174 eligible KPNC members, producing a 40% enrollment rate. Details are provided elsewhere,²⁴ but briefly, enrollment occurred an average of 2.3 (range=0.7-17.8) months after diagnosis and eligibility criteria included: female sex, 21 years or older, KPNC member, English, Spanish, Cantonese or Mandarin speaker, living within a 65-mile radius of a field interviewer, diagnosis of incident invasive breast cancer, and no prior history of other invasive cancers. Participants underwent an in-person baseline interview, and follow-up interviews by mail, phone, internet, or in person.

This study was approved by the Institutional Review Boards of KPNC and University of California, Berkeley. Written informed consent was obtained from all participants.

Dietary Assessment

Dietary intake was assessed at baseline and approximately 6, 24 and 72 months after, using a modified Block 2005 (NutritionQuest Questionnaires and Screeners, RRID:SCR_024429) Food Frequency Questionnaire (FFQ). The 139-item FFQ ascertained consumption of foods and beverages in the 6 months preceding each assessment. Three *a priori* diet quality indices (DQI) were calculated to measure concordance with dietary guidelines at each assessment: a healthy plant-based dietary index (hPDI),²⁵ an index based on the American Cancer Society cancer prevention guidelines (ACS),²⁶ and the 2015 Healthy Eating Index (HEI).²⁷ While ACS was selected because of its applicability to cancer outcomes, hPDI and HEI were chosen because of their prior demonstrated associations with breast cancer survival.^{28, 29} For each, higher scores indicated a more healthful dietary pattern (Table 1).

hPDI score

The hPDI score ranges from 17 to 85 points and is created from the sum of 17 dietary components, each combining foods that are similar in nutrient and culinary characteristics.²⁵ For 7 of these (vegetable oils, whole fruits, non-starchy vegetables, whole grains, legumes, nuts, and teas and coffees) greater consumption increases the hPDI score, while for the other 10 (dairy, animal fats, fruit juices, starchy vegetables, refined grains, eggs, seafood and fish, total meat, sweetened beverages and sweets), greater consumption reduces the score. Each component is worth up to 5 points, with assessment-specific quintiles of intake used for each dietary component.²⁵

ACS score

The ACS score ranges from 0 to 9 with three components: total fruits and vegetables (partly rewarding variety as consumption of 5 or more different fruits and vegetables per month), whole grains as a percent of total grains, and total red and processed meats. Each component is worth 0 (lowest) to 3 (highest) points and is based on cohort-specific quartiles for that component (the red and processed meat score is reversed).²⁶

2015 HEI Score

The HEI score, derived from the 2015-2020 Dietary Guidelines for Americans, contains 13 dietary subcomponents.²⁷ Six (total fruits, whole fruits, total vegetables, greens and beans, total protein foods and seafood and plant protein foods) are worth up to 5 points each, and seven others (whole grains, dairy, unsaturated fats, refined grains, sodium, added sugars and saturated fats) are worth up to 10 points each, for a possible total of 100 points. The HEI components are scored on a density basis (per 1,000 kcals or percentage of energy), except for unsaturated fats which is the ratio of unsaturated to saturated fats.²⁷

Physical Activity

Physical activity (PA) was assessed at baseline and at 6-, 24- and 72-months, using the Arizona Activity Questionnaire (AAQ).³⁰ The questionnaire assessed the frequency, duration, and intensity of activities across multiple domains, including work, recreation, household, and sedentary behaviors. The Compendium of Physical Activities,³¹ was used to assign a standard metabolic equivalent task value (MET) for each activity. A summary variable for hours per week of moderate to vigorous physical activity was created by multiplying each activity by frequency and duration, added over all activities at each follow-up with MET values of 3 or more.

Cigarette Smoking

The baseline questionnaire assessed the lifetime smoking history of each participant, and the 6-, 24- and 72-month questionnaires reassessed smoking status at each time point. Participants were asked if they currently or ever smoked cigarettes, and the average number of cigarettes per day smoked. Pack-years of smoking (SM) was calculated from the number of cigarettes per day divided by 20 (typical number of cigarettes per pack) and then multiplied by the number of years smoked.

Other Covariables

Demographic factors including age, race and ethnicity, and education were collected using a baseline questionnaire. Where possible, missing data were supplemented with data obtained from the KPNC electronic health records (EHR). Both body mass index (BMI) and comorbid conditions were obtained from the EHR. A weighted Elixhauser comorbidity index score was calculated from inpatient comorbid conditions dating back to the participant's earliest record.³² Diagnostic and clinical data, including tumor stage, estrogen receptor status (ER), progesterone receptor status (PR), human epidermal growth factor receptor 2 status (HER2), type of surgery, and receipt of chemotherapy, radiation therapy, and hormonal therapy were obtained from the KPNC Cancer Registry and other clinical databases.

All-Cause Mortality

While we recognize that all-cause mortality is a composite of multiple causes of death, it is in and of itself, an outcome with significant interest to both clinicians and patients. We therefore identified deaths from all causes through December 31st, 2019, primarily through linkage with KPNC mortality databases, the KPNC Cancer Registry, and linkages with data

from the State of California, the Social Security Administration, and the National Death Index. Deaths reported by relatives of participants were confirmed by medical chart review or KPNC mortality databases.

Final Analytic Sample

Of the 4,505 Pathways participants, 782 (17.4%) who did not complete the baseline FFQ were excluded from this analysis. An additional 63 (1.4%) participants were excluded due to extreme total energy intake at baseline (<400 or >4,000 kcal/day) and 2 (0.4%) more participants were excluded because they did not complete the baseline physical activity questionnaire. The final analytic cohort included 3,658 participants, of whom 734 (20.1%) died through the end of 2019 (Figure 1), of which, 34.7% died of breast cancer (Supplemental Table S1).

Final Analytic Dataset

The final analytic dataset contained 30 rows of data for each active participant during the entire 10-year study period (15 rows for the 5-year assessment). Each row represented a four-month time period in order to capture any dietary and lifestyle changes which may have occurred after baseline. For those participants who died or were censored during the study period, the number of rows for that participant was equal to the amount of time they were active, the last row representing their last 4-month time period. Using the baseline data as the start time of each covariate history, the study period for each participant began at the completion of their 6-month questionnaire with each covariate history lagging one follow-up period behind. Time-varying covariates with missing data were populated with values carried over from preceding time periods.

Hypothetical Interventions

For each of the DQIs and for physical activity, we specified 4 hypothetical interventions of increasing concordance with these measures, starting at baseline. For the dietary quality measures, we used the values associated with the 25th, 50th, 75th and 100th percentiles from the baseline distributions for each respective measure as the cut points for each of the hypothetical interventions. For example, the 25th percentile for the baseline distribution of hPDI was a score of 46, which then became the cut point for the first level of the hypothetical intervention for that index (Table 2). For physical activity, we used the recommendations from the Physical Activity Guidelines for Americans (PAGA)³³ as set forth by the US Department of Health and Human Services, as the basis for each cut point. The PAGA suggests that health benefits begin with as little as 1-hour per week from any type of physical activity, and substantial health benefits occur after 2.5 to 5-hours per week of moderate physical activity.³³ We therefore used 1, 2.5 and 5 as the cut points and added a fourth level at 10-hours per week since we included both moderate and vigorous physical activities.

For each hypothetical intervention, participants whose risk factor met the level of the intervention maintained that value, while those whose risk factor did not meet the hypothetical intervention level had their values set to that point. For example, for a hypothetical intervention on hPDI at the 46-point level, if a participant scored 46 or above

on their hPDI score they kept their observed value, whereas for those with hPDI values below 46, their hPDI score was set to 46. For the smoking intervention, smoking after baseline was set to the total pack-years of cigarettes smoked at baseline, corresponding to current smokers quitting smoking and non-smokers never starting. We additionally considered combinations of interventions on diet quality, physical activity, and smoking, as well as the standard course (no intervention). We specified all hypothetical interventions to be maintained over the entire 5- or 10-year follow-up period (Table 2).

Statistical Analysis

We estimated the 5- and 10-year risks of death under each intervention *via* parametric g-computation, a generalization of direct standardization, using a SAS macro developed by the Harvard Program on Causal Inference.³⁴ Briefly, g-computation fits a series of parametric regression models for the outcome and its predictors, then simulates these variables under specified values of exposures (Supplemental Figure S1).^{21, 23} Exposures and confounders are modeled sequentially accounting for the structural and temporal relationships between them as implied by a directed acyclic graph (DAG).³⁵ Fixing exposure levels can be interpreted similarly to an intervention that achieved that level of the variable. This allows estimation of the effects of a range of exposure comparisons on the outcome of interest (here, all-cause mortality), while correctly addressing time-varying confounding.²¹⁻²³ While the use of time-varying variables may be commonplace in standard epidemiological analyses, these studies are subject to the bias that arises from time-varying covariates, which both confound and mediate the exposure-outcome relationship.²²

For this analysis, a pooled logistic regression model for the outcome was fit across all time periods and regression models were fit to estimate the joint distribution of each covariate at each time interval, conditional on the covariate histories (Supplemental Table S2). Using Monte Carlo simulations, 3,658 covariate histories (equal to the sample size) consistent with the intervention were generated for each iteration. Using the estimated coefficients from the regression models, the values of each of the covariates were then calculated, and for the covariates that underwent the intervention, their values were then changed according to the intervention specification. When calculating the risk under the natural course (no intervention), no changes were made to any of the values of the predicted covariates. Average risk ratios (RR) and risk differences (RD) were calculated by comparing the population-level risk associated with each intervention to that associated with the natural course (no intervention). Corresponding 95% confidence intervals for these measures, were estimated *via* non-parametric bootstrapping on 1,000 samples.

To assess model specification, we compared the observed estimates with the model-based predicted estimates under no intervention, over the 10-year follow-up period for both the cumulative incidence for all-cause mortality (Figure 2) and each time-varying covariate (Supplemental Figure S2). Additionally, we conducted sensitivity analyses separately excluding BMI and alcohol intake, because even though the most current research supports their association with overall mortality among breast cancer survivors,^{7, 9} they have historically been considered equivocal risk factors.⁶ We further assessed model fit by applying quadratic and cubic functions as well as splines and interaction terms on BMI,

alcohol and other risk factors. There were no alternative model specifications that appeared to optimize model fit.

Fully adjusted models included time-varying data for hPDI, ACS, HEI, physical activity, and smoking. Independent models were assessed for each of the five primary exposures, however physical activity and smoking were included as covariables in models assessing hPDI, ACS and HEI. In models where physical activity and smoking were the primary exposures, HEI was included as a dietary covariable because of its intention to characterize a generally healthy diet. Models also included time-varying data on alcohol and energy intake, BMI and comorbidity scores. Additionally, the following baseline characteristics were included: age at diagnosis, race and ethnicity, education level, cancer stage, ER, PR, HER-2, surgery type, chemotherapy, radiation therapy, and hormonal therapy (variables categorized as specified in Table 3). Potential confounders of the relationship between diet, lifestyle factors and all-cause mortality were determined from existing knowledge and were incorporated into a DAG (Supplemental Figure S3). Those identified in the DAG as confounders were included in our final models.

Data availability

The data generated in this study are available upon request from the corresponding author.

RESULTS

Baseline Characteristics

A total of 3,658 participants were followed for 36,527 person-years. Average age at diagnosis was 59.7 years (range=24-94) and included whites (68.0%), Asians/Pacific-Islanders (13.0%), Hispanics (10.3%), Blacks (6.6%) and American Indians/Alaska Natives (2.1%). A complete description of demographic and clinical characteristics can be found in Table 3.

Hypothetical Interventions and All-Cause Mortality

There were 297 deaths within the first 5 years after a breast cancer diagnosis, 619 deaths within the first 10 years after a breast cancer diagnosis, and the Kaplan-Meier probabilities of death were 8.1% (95% CI 7.3-9.1) and 17.8% (95% CI 16.6-19.1), respectively. In comparison, the estimated 5- and 10-year risks of all-cause mortality under the natural course (no intervention), as calculated by the g-formula, were 7.9% (95% CI 7.1-8.9) and 17.4% (95% CI 16.0-18.8), respectively. Estimates of 5- and 10-year risks of death under each hypothetical intervention and the estimated RR and RD compared to the natural course are presented in Table 4. For individual hypothetical interventions on each DQI and physical activity, the inverse association with mortality became stronger with greater intensity of the intervention. For example, the 5-year RR comparing the hypothetical interventions on PA alone as 1, 2.5, 5 and 10 hours per week, compared to the natural course were 1.00 (95% CI 0.97-1.00), 0.96 (95% CI 0.92-0.99), 0.91 (95% CI 0.82-0.97) and 0.76 (95% CI 0.60-0.91), respectively. The same type of pattern was observed among the risk differences. The hypothetical intervention on smoking alone, whereby smokers quit smoking and non-smokers never started smoking, suggested a 6% decrease for both the 5-

and 10-year risk of all-cause mortality when compared to the natural course, though the estimate was not statistically significant in the case of the 5-year risk (5-year RR=0.94, 95% CI 0.88-1.00, 10-year RR=0.94, 95% CI 0.89-0.97).

Combining the moderate hypothetical interventions on diet quality and physical activity together with the intervention on smoking, suggested lower 5-year risks of mortality compared to the natural course (hPDI 46+PA 1+SM RR=0.89, 95% CI 0.82-0.94, ACS 3+PA 1+SM RR=0.91, 95% CI 0.85-0.96, HEI 66+PA 1+SM RR=0.91, 95% CI 0.84-0.95). Combining the most extreme hypothetical interventions, made the 5-year RRs stronger (hPDI=77+PA 10+SM RR=0.44, 95% CI 0.26-0.67, ACS=9+PA 10+SM RR=0.62, 95% CI 0.43-0.82, HEI=98+PA 10+SM RR=0.51, 95% CI 0.33-0.72). While the 10-year relative risks were slightly attenuated, the risk differences were more pronounced for both moderate and extreme hypothetical interventions (Table 4).

DISCUSSION

This study is the first to use a rigorous causal inference approach to demonstrate that higher diet quality, increased physical activity, and decreased smoking each could reduce mortality after breast cancer diagnosis. As suggested by the attenuating risk estimates produced from the hypothetical interventions on diet quality as compared to the natural course, increased intensity of the intervention on each dietary index (hPDI, ACS and HEI) was associated with lower risks of dying from all causes. Furthermore, increased intensity of the intervention on hPDI suggested the strongest reduction in risk, as compared to ACS and HEI. Like diet quality, the attenuating risk estimates associated with increased intensity of the hypothetical intervention on moderate to vigorous physical activity as compared to the natural course, suggested a reduction in risk of dying from all causes, as did the hypothetical intervention of stopping and never starting to smoke at or around the time of breast cancer diagnosis. Finally, the estimates associated with all three lifestyle behaviors (diet, physical activity and smoking) being jointly intervened upon at the same time, suggested that even modest interventions could provide the greatest reduction in risk.

To our knowledge, there have been no RCTs to jointly intervene on diet quality, physical activity, and smoking in order to assess their impact on all-cause mortality among breast cancer survivors. The existing evidence base consists primarily of observational studies that measure lifestyle factors at a single timepoint without accounting for factors that vary following a breast cancer diagnosis.⁷ To fill this gap, we estimated the dietary and lifestyle effects of hypothetical interventions with long-term follow-up on breast cancer survival, through the application of parametric g-computation. However, estimates obtained from previous studies should be compared to estimates from this analysis with caution, since such analyses target distinctly different estimands.^{21, 22} More specifically, estimates from this study represent the cumulative risk, or the probability of experiencing the event under a given hypothetical intervention, where estimates from prior observational studies using standard Cox regression models have produced conditional hazards under different covariate patterns.^{21, 22} Though our estimates may correspond more closely to the target measure under a clinical trial, we can still make qualitative comparisons.

Qualitatively speaking, our findings are consistent with other research on diet quality measured at or around the time of diagnosis and breast cancer survival. Four observational studies using *a priori* dietary quality indices (e.g. HEI, ACS),^{28, 36-38} and one observational study using data driven dietary patterns (e.g. unhealthy western diet vs healthy prudent diet),^{39, 40} each reported that increased concordance with healthful dietary patterns was associated with lower risk of all-cause mortality. To our knowledge, only two RCTs have examined a healthy dietary pattern among breast cancer survivors, the Women's Intervention Nutrition Study¹⁹ and the Women's Healthy Eating and Living Study,²⁰ in which the former focused on reduction of dietary fat and the latter on increased fruit and vegetable intake. While the results from both studies showed only modest effects on reducing all-cause mortality,^{19, 20} several recent systematic reviews and meta-analyses have reported on both pre- and post-diagnosis dietary patterns and all cause-mortality, each of which support a healthy dietary pattern to reduce the risk of all-cause mortality among breast cancer survival.¹¹⁻¹⁴

Our results suggest that increasing physical activity is inversely related with death after breast cancer, independent of diet quality or smoking. This is qualitatively supported by a recent prospective study showing patients meeting the highest level of Physical Activity Guidelines for Americans both before and 1 year after diagnosis had a 49% reduced risk of death, compared to those who met them the least.⁴¹ A sub-analysis, which also considered physical activity measured during treatment and 2 years post-diagnosis, revealed that the most active patients experienced a 69% reduced risk of death when compared to the least active.⁴¹ The overall relationship between pre- and postdiagnosis physical activity and reduced risk of all-cause mortality among breast cancer survivors is further supported by a recent meta-analysis that showed an 18% and 42% lower risk of all-cause mortality when comparing the highest versus lowest pre- and postdiagnosis physical activity categories, respectively.¹⁵

Our hypothetical intervention where smokers stopped and non-smokers did not start, showed a 6% reduction in risk of all-cause mortality compared to their natural course, for both 5 and 10-year follow-up periods (the 5-year RR was not statistically significant). This is qualitatively consistent with a meta-analysis from 2014, which included 9 studies, that reported a 33% increased risk of death when comparing those who smoked at the time of diagnosis to never smokers.⁴² Two studies prospectively examined post-diagnosis changes in smoking status and its impact on breast cancer survival, both of which supported an increased risk of death with continued smoking.^{18, 43}

Our study has several strengths, including a large sample drawn from a population of women newly diagnosed with invasive breast cancer, prospective longitudinal data collection with a long follow-up period, and repeated measures on dietary and lifestyle exposures and covariables. Using a causal framework allowed for evaluating several hypothetical individual and joint interventions that would otherwise be difficult to implement, and appropriately adjusted for known time-varying confounders that could lead to biased estimates in more traditional analytic approaches.

When there is no unmeasured confounding, no measurement error, no model misspecification, causal consistency, and all individuals have a nonzero probability of receiving all values of the treatment variable (positivity), g-computation can consistently estimate causal effects.^{21, 23} The g-formula relies heavily on a model that accurately represents the relationship between the exposure, the outcome, and any confounding variables. For this reason, we compared the observed estimates with the predicted model-based estimates under no intervention for our outcome and for all time-varying covariates to assess model fit. Most notable of these comparisons, were the near identical curves for both the observed and predicted cumulative incidence estimates for all-cause mortality (Figure 2). This, along with the absence of any egregious differences between the observed and predicted means among the time-varying covariates (Supplemental Figure S2) increased confidence in our model specification. Finally, as a sensitivity exercise, we reversed the order in which the g-formula modelled the time-varying covariates and found differences in the estimates to be negligible.

Despite the strengths of our study, some limitations should be acknowledged. Although g-computation can properly account for known confounders, it does not resolve bias due to unknown or unmeasured sources of confounding. We adjusted for a number of important covariates, informed by a causal diagram based on expert knowledge and literature review. However, we acknowledge that there may be important factors that we were unable to account for, which would result in a violation of the exchangeability assumption required for causal interpretation.

A key requirement for causal interpretation of our estimates is that the exposures reflect the same experience for every individual, often referred to as “causal consistency.” The extent that a violation of this assumption hinders meaningful inference has been the source of controversy⁴⁴⁻⁴⁷ but the general implication is that some “versions” of the exposure may yield different outcomes than others (e.g., achieving a specific HEI score by increasing vegetable intake may influence mortality differently than limiting processed meat consumption). We acknowledge that the consistency assumption is likely to be violated to some degree for lifestyle factors such as diet and physical activity, yet there are strong arguments that even with such violations, these estimates represent more relevant parameters for clinical and public health investigations.^{44, 47}

G-computation is also subject to the ‘g-null paradox’, a phenomenon in which it is not possible to correctly specify the analytic model, and as a result, the null hypothesis is rejected, even if true.²¹ As described in Taubman et al (2009)²¹, to mitigate this issue, we only considered primary exposures that we *a priori* believed were related to the outcome, so that any observed associations would unlikely be explained by the ‘g-null paradox’. Although we specified the models involved in the g-computation to the best of our ability, it is possible that model misspecification could have influenced our results. In addition, immortal person-time, or the length of time between diagnosis and study enrollment, may be a source of bias, but methods to account for it in this setting are not well developed. However, with over 96% of participants having completed the baseline questionnaire within 120 days of diagnosis, we do not believe that bias due to immortal person-time to be significant. To further confirm, we performed a sensitivity analysis comparing the risk

associated with the natural course after excluding all participants recruited more than 120 days after their breast cancer diagnosis with the risk associated with the natural course when including everyone, and the differences were negligible.

G-computation also does not directly address limitations arising from measurement error of lifestyle factors, especially dietary intakes.⁴⁸ This could result in biased estimates and/or misclassification, perhaps under or over estimating the effects of the hypothetical interventions. However, by using research instruments such as the FFQ and AAQ, which have been repeatedly validated in different settings and shown to be effective measurement tools,³⁰ we do not believe misclassification of diet and physical activity to be significant. Epidemiologic research has also shown that adult retrospective self-reports of average daily cigarette use to be a relatively accurate means of estimating average number of cigarettes smoked.⁴⁹ In addition, all hypothetical interventions are assumed to be maintained over the entire 5- or 10-year follow-up period, which does not account for potential fluctuations in adherence. Furthermore, we recognize that dietary index scores, in general, are limited by the subjectivity of their specified food groups, the dependence on their underlying guidelines, and the potential for classifying individuals into broad categories that may not reveal the nuances of a healthy diet.⁵⁰ It is also difficult to interpret the results from dietary index scores beyond the study population in which they were examined, since there are no specific recommendations that accompany the scores. However, testing the scores at varying levels of the index, as we have done in this study, provides insight into the underlying effect. We also recognize that in a real-world setting, it may be difficult to identify a cohort of breast cancer survivors who could adhere to a diet quality and lifestyle intervention at the highest levels over long follow-up periods (as illustrated in some of our analyses here), however we felt that it was important to describe the estimates from our models along a range, from modest to most extreme, to help support more informed decision making. What may seem like an unachievable intervention in practice, may be achievable for certain subsets of the population since the hypothetical interventions were based on observed values.

Finally, it is also possible that women with breast cancer who enrolled in this study might be systematically different from those who did not. However, other data shows that the Pathways Study cohort reflects the eligible population, with minor shifts toward slightly younger ages, higher proportion of whites and earlier stages at diagnosis (Supplemental Table S3). Moreover, one study showed that the 1996 through 2009 KPNC population of women diagnosed with breast cancer represented one-third of the breast cancer population in Northern California, as compared to non-KPNC cancer centers which represented only 7%, and all other non-KPNC hospitals which represented 61%.⁵¹ Additionally, breast cancer patients from KPNC were more demographically and clinically similar to breast cancer patients from the non-KPNC hospitals than they were to breast cancer patients from non-KPNC cancer centers, suggesting that KPNC is more likely to be representative of the general underlying population of breast cancer survivors than those from the academic medical centers where much of clinical research is currently conducted.⁵¹

In summary, interventions that increase diet quality, increase physical activity, or eliminate/prevent smoking, could reduce deaths among breast cancer survivors. Additionally, even

modest interventions on all three risk factors simultaneously could provide the greatest benefit, at both 5 and 10 years out from diagnosis. Future research may include the use of similar methodological approaches to examining the impact of diet and lifestyle factors on other breast cancer survival endpoints, such as recurrence and second primary cancers. These findings could also be further validated by implementing intervention trials among targeted subgroups who are motivated to modify their dietary and lifestyle behavior. The results from this study highlight the lifesaving potential of diet and lifestyle interventions on breast cancer survivors at the time of their diagnosis and add new rigor to the evidence base for lifestyle guidelines for women with breast cancer.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

ACKNOWLEDGEMENTS

We gratefully acknowledge the Pathways Study participants for their vital contributions to this research.

Funding

Elizabeth M. Feliciano: American Institute for Cancer Research (632996), National Cancer Institute at the National Institutes of Health (K01-CA226155).

Lawrence H. Kushi: National Cancer Institute at the National Institutes of Health (CA-195565, R01-CA105274).

Abbreviations

AAQ	Arizona Activity Questionnaire
ACS	American Cancer Society
CI	Confidence Interval
DAG	Directed Acyclic Graph
DQI	Diet Quality Index
EHR	Electronic Health Record
ER	Estrogen Receptor
FFQ	Food Frequency Questionnaire
HEI	Healthy Eating Index
HER2	Human Epidermal Growth Factor Receptor 2
hPDI	healthy Plant-Based Dietary Index
KPNC	Kaiser Permanente Northern California
MET	Metabolic Equivalent Task
PA	Physical Activity

PAGA	Physical Activity Guidelines for Americans
PR	Progesterone Receptor
RCT	Randomized Controlled Trial
RD	Risk Difference
RR	Risk Ratio
SM	Smoking

REFERENCES

1. Breast cancer statistics. World Cancer Reserach Fund; 2022 [cited 2022 09/20/2022]; Available from: <https://www.wcrf.org/dietandcancer/breast-cancer-statistics/>.
2. Key Statistics for Breast Cancer. American Cancer Society; 2022 [cited 2022 09/20/2022]; Available from: <https://www.cancer.org/cancer/breast-cancer/about/how-common-is-breast-cancer.html>.
3. Salminen EK, Lagstrom HK, Heikkila S, Salminen S. Does breast cancer change patients' dietary habits? *Eur J Clin Nutr.* 2000; 54:844–8. [PubMed: 11114679]
4. Salminen E, Heikkila S, Poussa T, Lagstrom H, Saario R, Salminen S. Female patients tend to alter their diet following the diagnosis of rheumatoid arthritis and breast cancer. *Prev Med.* 2002; 34:529–35. [PubMed: 11969354]
5. McBride CM, Ostroff JS. Teachable moments for promoting smoking cessation: the context of cancer care and survivorship. *Cancer Control.* 2003; 10:325–33. [PubMed: 12915811]
6. Diet, Nutrition, Physical Activity and Cancer: a Global Perspective: World Cancer Research Fund/ American Institute for Cancer Research 2018.
7. Rock CL, Thomson CA, Sullivan KR, Howe CL, Kushi LH, Caan BJ, et al. American Cancer Society nutrition and physical activity guideline for cancer survivors. *CA Cancer J Clin.* 2022; 72:230–62. [PubMed: 35294043]
8. Duan W, Li S, Meng X, Sun Y, Jia C. Smoking and survival of breast cancer patients: A meta-analysis of cohort studies. *Breast.* 2017; 33:117–24. [PubMed: 28371644]
9. Patterson RE, Cadmus LA, Emond JA, Pierce JP. Physical activity, diet, adiposity and female breast cancer prognosis: a review of the epidemiologic literature. *Maturitas.* 2010; 66:5–15. [PubMed: 20097494]
10. Cespedes EM, Hu FB. Dietary patterns: from nutritional epidemiologic analysis to national guidelines. *The American journal of clinical nutrition.* 2015; 101:899–900. [PubMed: 25832336]
11. Jochems SHJ, Van Osch FHM, Bryan RT, Wesselius A, van Schooten FJ, Cheng KK, et al. Impact of dietary patterns and the main food groups on mortality and recurrence in cancer survivors: a systematic review of current epidemiological literature. *BMJ Open.* 2018; 8:e014530.
12. Terranova CO, Protani MM, Reeves MM. Overall Dietary Intake and Prognosis after Breast Cancer: A Systematic Review. *Nutrition and cancer.* 2018; 70:153–63. [PubMed: 29308928]
13. Molina-Montes E, Salamanca-Fernandez E, Garcia-Villanova B, Sanchez MJ. The Impact of Plant-Based Dietary Patterns on Cancer-Related Outcomes: A Rapid Review and Meta-Analysis. *Nutrients.* 2020; 12.
14. Schwedhelm C, Boeing H, Hoffmann G, Aleksandrova K, Schwingshackl L. Effect of diet on mortality and cancer recurrence among cancer survivors: a systematic review and meta-analysis of cohort studies. *Nutr Rev.* 2016; 74:737–48. [PubMed: 27864535]
15. Friedenreich CM, Stone CR, Cheung WY, Hayes SC. Physical Activity and Mortality in Cancer Survivors: A Systematic Review and Meta-Analysis. *JNCI Cancer Spectr.* 2020; 4:pkz080. [PubMed: 32337494]
16. Spei ME, Samoli E, Bravi F, La Vecchia C, Bamia C, Benetou V. Physical activity in breast cancer survivors: A systematic review and meta-analysis on overall and breast cancer survival. *Breast.* 2019; 44:144–52. [PubMed: 30780085]

17. Morishita S, Hamaue Y, Fukushima T, Tanaka T, Fu JB, Nakano J. Effect of Exercise on Mortality and Recurrence in Patients With Cancer: A Systematic Review and Meta-Analysis. *Integr Cancer Ther.* 2020; 19:1534735420917462.
18. Parada H Jr., Bradshaw PT, Steck SE, Engel LS, Conway K, Teitelbaum SL, et al. Postdiagnosis Changes in Cigarette Smoking and Survival Following Breast Cancer. *JNCI Cancer Spectr.* 2017; 1.
19. Pierce JP, Natarajan L, Caan BJ, Parker BA, Greenberg ER, Flatt SW, et al. Influence of a diet very high in vegetables, fruit, and fiber and low in fat on prognosis following treatment for breast cancer: the Women's Healthy Eating and Living (WHEL) randomized trial. *JAMA.* 2007; 298:289–98. [PubMed: 17635889]
20. Chlebowski RT, Blackburn GL, Thomson CA, Nixon DW, Shapiro A, Hoy MK, et al. Dietary fat reduction and breast cancer outcome: interim efficacy results from the Women's Intervention Nutrition Study. *Journal of the National Cancer Institute.* 2006; 98:1767–76. [PubMed: 17179478]
21. Taubman SL, Robins JM, Mittleman MA, Hernan MA. Intervening on risk factors for coronary heart disease: an application of the parametric g-formula. *Int J Epidemiol.* 2009; 38:1599–611. [PubMed: 19389875]
22. Keil AP, Edwards JK, Richardson DB, Naimi AI, Cole SR. The parametric g-formula for time-to-event data: intuition and a worked example. *Epidemiology.* 2014; 25:889–97. [PubMed: 25140837]
23. Lajous M, Willett WC, Robins J, Young JG, Rimm E, Mozaffarian D, et al. Changes in fish consumption in midlife and the risk of coronary heart disease in men and women. *Am J Epidemiol.* 2013; 178:382–91. [PubMed: 23813701]
24. Kwan ML, Ambrosone CB, Lee MM, Barlow J, Krathwohl SE, Ergas IJ, et al. The Pathways Study: a prospective study of breast cancer survivorship within Kaiser Permanente Northern California. *Cancer causes & control : CCC.* 2008; 19:1065–76. [PubMed: 18478338]
25. Satija A, Bhupathiraju SN, Spiegelman D, Chiuve SE, Manson JE, Willett W, et al. Healthful and Unhealthful Plant-Based Diets and the Risk of Coronary Heart Disease in U.S. Adults. *J Am Coll Cardiol.* 2017; 70:411–22. [PubMed: 28728684]
26. Kushi LH, Doyle C, McCullough M, Rock CL, Demark-Wahnefried W, Bandera EV, et al. American Cancer Society Guidelines on nutrition and physical activity for cancer prevention: reducing the risk of cancer with healthy food choices and physical activity. *CA Cancer J Clin.* 2012; 62:30–67. [PubMed: 22237782]
27. Krebs-Smith SM, Pannucci TE, Subar AF, Kirkpatrick SI, Lerman JL, Tooze JA, et al. Update of the Healthy Eating Index: HEI-2015. *J Acad Nutr Diet.* 2018; 118:1591–602. [PubMed: 30146071]
28. Ergas IJ, Cespedes Feliciano EM, Bradshaw PT, Roh JM, Kwan ML, Cadenhead J, et al. Diet Quality and Breast Cancer Recurrence and Survival: The Pathways Study. *JNCI Cancer Spectr.* 2021; 5:pkab019. [PubMed: 33928215]
29. Anyene IC, Ergas IJ, Kwan ML, Roh JM, Ambrosone CB, Kushi LH, et al. Plant-Based Dietary Patterns and Breast Cancer Recurrence and Survival in the Pathways Study. *Nutrients.* 2021; 13:3374. [PubMed: 34684375]
30. Staten LK, Taren DL, Howell WH, Tobar M, Poehlman ET, Hill A, et al. Validation of the Arizona Activity Frequency Questionnaire using doubly labeled water. *Med Sci Sports Exerc.* 2001; 33:1959–67. [PubMed: 11689750]
31. Ainsworth BE, Haskell WL, Herrmann SD, Meckes N, Bassett DR Jr., Tudor-Locke C, et al. 2011 Compendium of Physical Activities: a second update of codes and MET values. *Med Sci Sports Exerc.* 2011; 43:1575–81. [PubMed: 21681120]
32. van Walraven C, Austin PC, Jennings A, Quan H, Forster AJ. A modification of the Elixhauser comorbidity measures into a point system for hospital death using administrative data. *Med Care.* 2009; 47:626–33. [PubMed: 19433995]
33. Physical Activity Guidelines for Americans. Centers for Disease Control and Prevention; [cited 2022]; Available from: https://health.gov/sites/default/files/2019-09/Physical_Activity_Guidelines_2nd_edition.pdf.

34. GFORMULA 3.0 - The parametric g-formula in SAS. Harvard T.H. Chan Casaulab; [cited 2022 09/20/2022]; Available from: <https://causalab.sph.harvard.edu/software/>.
35. Greenland S, Pearl J, Robins JM. Causal diagrams for epidemiologic research. *Epidemiology*. 1999; 10:37–48. [PubMed: 9888278]
36. George SM, Ballard-Barbash R, Shikany JM, Caan BJ, Freudenheim JL, Kroenke CH, et al. Better postdiagnosis diet quality is associated with reduced risk of death among postmenopausal women with invasive breast cancer in the women's health initiative. *Cancer Epidemiol Biomarkers Prev*. 2014; 23:575–83. [PubMed: 24493629]
37. George SM, Irwin ML, Smith AW, Neuhouser ML, Reedy J, McTiernan A, et al. Postdiagnosis diet quality, the combination of diet quality and recreational physical activity, and prognosis after early-stage breast cancer. *Cancer causes & control : CCC*. 2011; 22:589–98. [PubMed: 21340493]
38. Wang F, Cai H, Gu K, Shi L, Yu D, Zhang M, et al. Adherence to Dietary Recommendations among Long-Term Breast Cancer Survivors and Cancer Outcome Associations. *Cancer Epidemiol Biomarkers Prev*. 2020; 29:386–95. [PubMed: 31871105]
39. Kroenke CH, Fung TT, Hu FB, Holmes MD. Dietary patterns and survival after breast cancer diagnosis. *J Clin Oncol*. 2005; 23:9295–303. [PubMed: 16361628]
40. Kwan ML, Weltzien E, Kushi LH, Castillo A, Slattery ML, Caan BJ. Dietary patterns and breast cancer recurrence and survival among women with early-stage breast cancer. *J Clin Oncol*. 2009; 27:919–26. [PubMed: 19114692]
41. Cannioto RA, Hutson A, Dighe S, McCann W, McCann SE, Zirpoli GR, et al. Physical Activity Before, During, and After Chemotherapy for High-Risk Breast Cancer: Relationships With Survival. *Journal of the National Cancer Institute*. 2021; 113:54–63. [PubMed: 32239145]
42. Berube S, Lemieux J, Moore L, Maunsell E, Brisson J. Smoking at time of diagnosis and breast cancer-specific survival: new findings and systematic review with meta-analysis. *Breast Cancer Res*. 2014; 16:R42. [PubMed: 24745601]
43. Passarelli MN, Newcomb PA, Hampton JM, Trentham-Dietz A, Titus LJ, Egan KM, et al. Cigarette Smoking Before and After Breast Cancer Diagnosis: Mortality From Breast Cancer and Smoking-Related Diseases. *J Clin Oncol*. 2016; 34:1315–22. [PubMed: 26811527]
44. van der Laan MJ, Haight TJ, Tager IB. van der Laan et al. Respond to “Hypothetical Interventions to Define Causal Effects”. *American Journal of Epidemiology*. 2005; 162:621–2.
45. Hernan MA, Taubman SL. Does obesity shorten life? The importance of well-defined interventions to answer causal questions. *Int J Obes (Lond)*. 2008; 32 Suppl 3:S8–14.
46. Petersen ML. Compound treatments, transportability, and the structural causal model: the power and simplicity of causal graphs. *Epidemiology*. 2011; 22:378–81. [PubMed: 21464653]
47. Pearl J. On the consistency rule in causal inference: axiom, definition, assumption, or theorem? *Epidemiology*. 2010; 21:872–5. [PubMed: 20864888]
48. Willett W. Nutritional epidemiology: issues and challenges. *Int J Epidemiol*. 1987; 16:312–7. [PubMed: 3610460]
49. Blank MD, Breland AB, Enlow PT, Duncan C, Metzger A, Cobb CO. Measurement of smoking behavior: Comparison of self-reports, returned cigarette butts, and toxicant levels. *Exp Clin Psychopharmacol*. 2016; 24:348–55. [PubMed: 27347741]
50. Moeller SM, Reedy J, Millen AE, Dixon LB, Newby PK, Tucker KL, et al. Dietary patterns: challenges and opportunities in dietary patterns research an Experimental Biology workshop, April 1, 2006. *J Am Diet Assoc*. 2007; 107:1233–9. [PubMed: 17604756]
51. Gomez SL, Shariff-Marco S, Von Behren J, Kwan ML, Kroenke CH, Keegan TH, et al. Representativeness of breast cancer cases in an integrated health care delivery system. *BMC Cancer*. 2015; 15:688. [PubMed: 26467773]

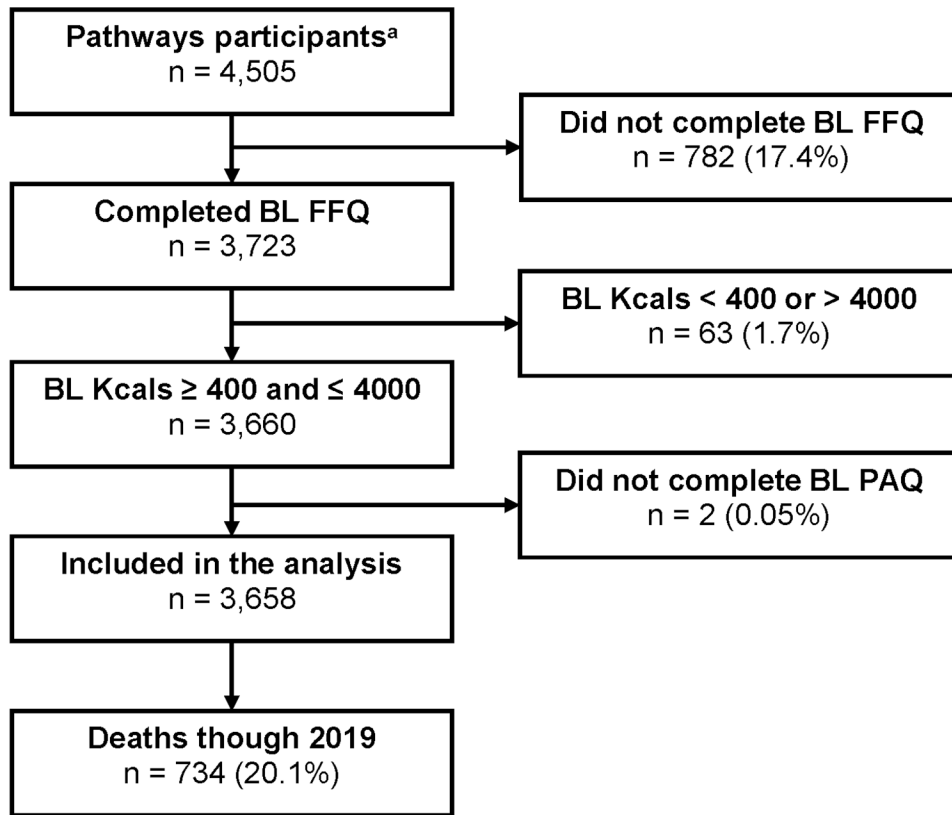


Figure 1.

Analytic exclusions and deaths among participants from the Pathways Study. This flowchart depicts the order and type of exclusions made leading up to the final analytic cohort, and therewithin, the total number of deaths.

Abbreviations: BL (baseline), FFQ (food frequency questionnaire), Kcals (kilocalories), PAQ (physical activity questionnaire).

^aEnrolled at or around the time of their first invasive breast cancer diagnosis, 2005-2013.

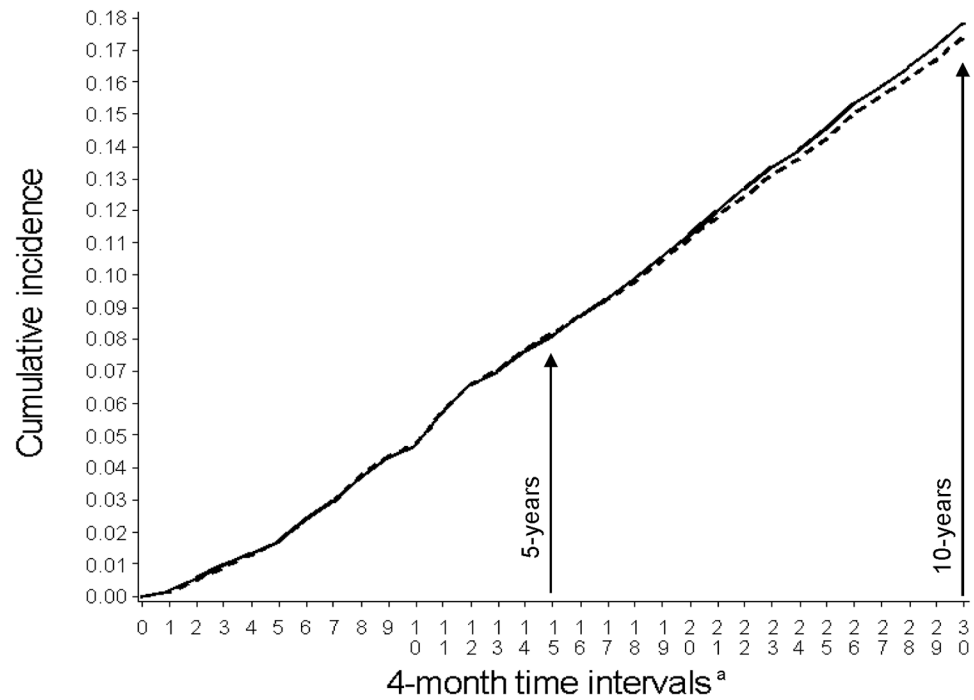


Figure 2. Cumulative incidence for all-cause mortality, comparing the observed (solid line) vs natural course (dotted line) estimates by follow-up in 4-month time intervals. This figure compares the observed estimates with the model-based predicted estimates under no intervention using the g-formula.

^aEach interval represents a 4-month time period.

Table 1.

Scoring methods and optimal quantities for each component of hPDI, ACS and HEI

Food group	hPDI (max = 85 points)	ACS (max = 9 points)	HEI^a (max = 100 points)
Dairy	= Lowest quintile		1.3 cups/1,000 kcal
Fats			
Animal fats	= Lowest quintile		
Saturated fats			8% of energy
Unsaturated fats			PUFAs + MUFAs:SFA 2.5
Vegetable oils	= Highest quintile		
Fruits and vegetables			
Fruits			
Fruit juices	= Lowest quintile		
Whole fruits	= Highest quintile		0.4 cups/1,000 kcal
Total fruit			0.8 cups/1,000 kcal
Vegetables			
Greens and beans			0.2 cups/1,000 kcal
Non-Starchy vegetables	= Highest quintile ^b		
Starchy vegetables	= Lowest quintile ^c		
Total vegetables			1.1 cups/1,000 kcal
Total fruits and vegetables		= Highest quartile ^d	
Grains			
Refined grains	= Lowest quintile		1.8 ounces/1,000 kcal
Whole grains	= Highest quintile	= Highest quartile ^e	> 1.5 ounces/1,000 kcal
Proteins			
Eggs	= Lowest quintile		
Legumes	= Highest quintile ^f		
Nuts	= Highest quintile		
Red and processed meats		= Lowest quartile	
Seafood and fish	= Lowest quintile		
Seafood and plant proteins			0.8 ounces/1,000 kcal
Total Meat	= Lowest quintile ^g		
Total protein			2.5 ounces/1,000 kcal
Sodium			1.1 grams/1,000 kcal
Sugar			
Added sugar			6.5% of energy
Sweetened beverages	= Lowest quintile		
Sweets	= Lowest quintile		
Teas and coffees	= Highest quintile		

Abbreviations: hPDI (healthy plant-based diet score), ACS (American Cancer Society nutrition guidelines score), HEI (2015 Healthy Eating Index score), PUFA (polyunsaturated fat), MUFA (monounsaturated fat), SFA (saturated fat).

^aAll quantities are average amounts per day.

^bDoes not include potatoes and legumes.

^cDoes not include sweet potatoes which are included in non-starchy vegetables.

^dExcludes fruit juices and potatoes. Includes partial variety score for consumption of 5 different fruits or vegetables per month.

^eWhole grains calculated as a percent of total grains.

^fLegumes include beans, peas, soybeans and soy products such as tofu.

^gIncludes red and processed meats, poultry and organ meat.

Author Manuscript

Author Manuscript

Author Manuscript

Author Manuscript

Table 2.Descriptions of hypothetical interventions at baseline for Pathways Study participants^a

Intervention	Description
(1) Natural course	Observed data: no intervention on treatment at any time.
Diet quality ^b	
(2) hPDI 46	Any hPDI score < 46 is set to 46 and any hPDI score 46 is left unchanged.
(3) hPDI 51	Any hPDI score < 51 is set to 51 and any hPDI score 51 is left unchanged.
(4) hPDI 57	Any hPDI score < 57 is set to 57 and any hPDI score 57 is left unchanged.
(5) hPDI = 77	Any hPDI score < 77 is set to 77 and any hPDI score 77 is left unchanged.
(6) ACS 3	Any ACS score < 3 is set to 3 and any ACS 3 is left unchanged.
(7) ACS 4	Any ACS score < 4 is set to 4 and any ACS 4 is left unchanged.
(8) ACS 6	Any ACS score < 6 is set to 6 and any ACS 6 is left unchanged.
(9) ACS = 9	Any ACS score < 9 is set to 9 and any ACS 9 is left unchanged.
(10) HEI 66	Any HEI score < 66 is set to 66 and any HEI score 66 is left unchanged.
(11) HEI 73	Any HEI score < 73 is set to 73 and any HEI score 73 is left unchanged.
(12) HEI 79	Any HEI score < 79 is set to 79 and any HEI score 79 is left unchanged.
(13) HEI = 98	Any HEI score < 98 is set to 98 and any HEI score 98 is left unchanged.
Physical activity ^c	
(14) PA 1	Any PA score < 1 is set to 1 and any PA score 1 is left unchanged.
(15) PA 2.5	Any PA score < 2.5 is set to 2.5 and any PA score 2.5 is left unchanged.
(16) PA 5	Any PA score < 5 is set to 5 and any PA score 5 is left unchanged.
(17) PA 10	Any PA score < 10 is set to 10 and any PA score 10 is left unchanged.
Smoking	
(18) Quit/Never start	Smokers quit at baseline and non-smokers never start.
Combined	
(20) 2 + 14 + 18	Joint intervention of hPDI 48 and PA 1 and No Smoking.
(21) 5 + 17 + 18	Joint intervention of hPDI = 81 and PA 10 and No Smoking.
(22) 6 + 14 + 18	Joint intervention of ACS 3 and PA 1 and No Smoking.
(23) 9 + 17 + 18	Joint intervention of ACS = 9 and PA 10 and No Smoking.
(24) 10 + 14 + 18	Joint intervention of HEI 66 and PA 1 and No Smoking.
(25) 13 + 17 + 18	Joint intervention of HEI = 98 and PA 10 and No Smoking.

Abbreviations: hPDI (healthy plant-based dietary index score), ACS (American Cancer Society nutrition guidelines score), HEI (Healthy Eating Index score), PA (moderate to vigorous physical activity in hr/week)

^aAll hypothetical interventions are assumed to be maintained over the entire 5- or 10-year follow-up period.

^bDiet quality intervention cut points determined from the 25th, 50th, 75th and 100th percentiles of the baseline distributions for each dietary index.

^cAll Physical activity intervention cut points determined from Physical Activity Guidelines for Americans as set forth by the United States Department of Health and Human Services (see text).

Table 3.

Baseline characteristics of the Pathways Study participants (n=3,658)

Demographic and behavioral characteristics	No. (%)	Mean (SD)
Age at diagnosis (years)		59.7 (11.9)
Race/ethnicity		
White	2,489 (68.0)	
Black	240 (6.6)	
Asian/Pacific Islander	475 (13.0)	
Hispanic	378 (10.3)	
American Indian/Alaska Native	76 (2.1)	
Education		
High school or less	547 (15.0)	
Some college	1,245 (34.0)	
College graduate	1,024 (28.0)	
Postgraduate	842 (23.0)	
BMI (kg/m ²)		28.4 (6.7)
Physical activity (hours per week) ^a		5.9 (5.5)
Smoking (total pack-years) ^b		7.4 (15.6)
Alcohol intake (grams/d)		7.1 (12.8)
Energy intake (kcal/d)		1,466.1 (568.5)
Elixhauser comorbidity score		0.2 (2.7)
hPDI score		51.2 (7.6)
ACS score		4.4 (2.1)
HEI score		72.0 (9.5)
Clinical characteristics		
Cancer stage		
I	2,006 (54.8)	
II	1,250 (34.2)	
III	346 (9.5)	
IV	56 (1.5)	
ER status		
Positive	3,070 (83.9)	
Negative	586 (16.0)	
Unknown	2 (0.1)	
PR status		
Positive	2,346 (64.1)	
Negative	1,307 (35.7)	
Unknown	5 (0.1)	
HER2 status		
Positive	471 (12.9)	
Negative	3,044 (83.2)	
Unknown	143 (3.9)	

Surgery type	
Lumpectomy	2,175 (59.5)
Mastectomy	1,364 (37.3)
None	117 (3.2)
Unknown	2 (0.1)
Chemotherapy	
No	1,939 (53.0)
Yes	1,708 (46.7)
Unknown	11 (0.3)
Radiation therapy	
No	2,038 (55.7)
Yes	1,620 (44.3)
Hormonal therapy	
No	902 (24.7)
Yes	2,731 (74.7)
Unknown	25 (0.7)

Abbreviations: SD (standard deviation), BMI (body mass index), hPDI (healthy plant-based dietary index), ACS (American Cancer Society nutrition guidelines), HEI (Healthy Eating Index), ER (estrogen receptor), PR (progesterone receptor), HER2 (human epidermal growth factor receptor 2).

^aIncludes moderate to vigorous physical activities. There were 3 participants with unknown values.

^bThere were 21 participants with unknown values.

Estimates of 5- and 10-year risks and 95% confidence intervals from hypothetical interventions on diet and lifestyle and all-cause mortality^a

Table 4.

Intervention	5-Year Risk (No. deaths = 297)			10-Year Risk (No. deaths = 619)		
	Risk (95% CI) ^b	RR (95% CI)	RD (95% CI) ^b	Risk (95% CI) ^b	RR (95% CI)	RD (95% CI) ^b
(1) Natural course	7.9 (7.1, 8.9)	1	0	17.4 (16.0, 18.8)	1	0
Diet quality ^{c,d}						
(2) hPDI 46	7.7 (6.8, 8.6)	0.97 (0.93, 1.00)	-0.25 (-0.58, -0.01)	17.0 (15.7, 18.6)	0.97 (0.96, 1.01)	-0.46 (-0.73, 0.19)
(3) hPDI 51	7.4 (6.4, 8.4)	0.93 (0.85, 0.99)	-0.57 (-1.18, -0.06)	16.9 (15.3, 18.4)	0.97 (0.92, 1.02)	-0.53 (-1.40, 0.29)
(4) hPDI 57	6.8 (5.7, 8.1)	0.86 (0.74, 0.99)	-1.13 (-2.15, -0.11)	16.3 (14.4, 18.3)	0.93 (0.85, 1.02)	-1.14 (-2.51, 0.40)
(5) hPDI = 77	5.1 (3.1, 7.6)	0.64 (0.40, 0.94)	-2.87 (-4.87, -0.43)	14.4 (10.6, 18.6)	0.83 (0.62, 1.07)	-3.00 (-6.55, 1.15)
(6) ACS 3	7.9 (6.9, 8.8)	0.99 (0.96, 1.02)	-0.05 (-0.34, 0.14)	17.3 (15.8, 18.6)	0.99 (0.96, 1.01)	-0.18 (-0.66, 0.21)
(7) ACS 4	7.8 (6.8, 8.8)	0.99 (0.92, 1.03)	-0.10 (-0.60, 0.24)	16.7 (15.6, 18.5)	0.96 (0.94, 1.01)	-0.69 (-1.09, 0.24)
(8) ACS 6	7.5 (6.3, 9.0)	0.95 (0.83, 1.07)	-0.40 (-1.37, 0.55)	16.3 (14.6, 18.4)	0.93 (0.86, 1.03)	-1.15 (-2.37, 0.49)
(9) ACS = 9	7.1 (5.3, 9.4)	0.90 (0.69, 1.15)	-0.82 (-2.53, 1.19)	15.6 (12.9, 18.7)	0.89 (0.75, 1.06)	-1.87 (-4.37, 1.01)
(10) HEI 66	7.7 (6.8, 8.6)	0.97 (0.94, 1.00)	-0.25 (-0.52, 0.00)	17.1 (15.8, 18.6)	0.98 (0.97, 1.02)	-0.31 (-0.58, 0.27)
(11) HEI 73	7.4 (6.4, 8.4)	0.94 (0.86, 1.00)	-0.51 (-1.07, -0.01)	17.0 (15.5, 18.6)	0.97 (0.94, 1.02)	-0.45 (-1.14, 0.39)
(12) HEI 79	7.0 (5.8, 8.3)	0.89 (0.78, 0.99)	-0.90 (-1.73, -0.06)	16.6 (15.0, 18.6)	0.95 (0.89, 1.04)	-0.90 (-1.88, 0.62)
(13) HEI = 98	5.7 (4.0, 8.0)	0.72 (0.53, 0.99)	-2.22 (-3.73, -0.11)	15.5 (12.6, 19.2)	0.89 (0.74, 1.09)	-1.90 (-4.54, 1.50)
Physical activity ^{e,f}						
(14) PA 1	7.9 (7.0, 8.8)	1.00 (0.97, 1.00)	-0.04 (-0.26, 0.03)	17.1 (15.8, 18.6)	0.98 (0.97, 1.00)	-0.37 (-0.56, 0.07)
(15) PA 2.5	7.6 (6.7, 8.5)	0.96 (0.92, 0.99)	-0.28 (-0.65, -0.07)	16.6 (15.2, 18.2)	0.95 (0.93, 0.98)	-0.87 (-1.21, -0.30)
(16) PA 5	7.2 (6.1, 8.2)	0.91 (0.82, 0.97)	-0.74 (-1.47, -0.27)	15.5 (14.0, 17.3)	0.89 (0.84, 0.95)	-1.90 (-2.68, -0.92)
(17) PA 10	6.0 (4.6, 7.5)	0.76 (0.60, 0.91)	-1.91 (-3.17, -0.70)	13.1 (11.0, 15.5)	0.75 (0.66, 0.87)	-4.31 (-6.03, -2.30)
(18) No Smoking ^g	7.5 (6.5, 8.5)	0.94 (0.88, 1.00)	-0.47 (-0.93, -0.04)	16.3 (14.7, 17.9)	0.94 (0.89, 0.97)	-1.10 (-1.85, -0.54)
(20) 2 + 14 + 18	7.1 (6.1, 8.0)	0.89 (0.82, 0.94)	-0.86 (-1.43, -0.45)	15.9 (14.2, 17.2)	0.91 (0.86, 0.95)	-1.54 (-2.42, -0.93)
(21) 5 + 17 + 18	3.5 (2.0, 5.4)	0.44 (0.26, 0.67)	-4.45 (-5.97, -2.57)	10.2 (7.0, 13.8)	0.58 (0.41, 0.78)	-7.29 (-10.11, -3.77)
(22) 6 + 14 + 18	7.2 (6.2, 8.2)	0.91 (0.85, 0.96)	-0.72 (-1.24, -0.30)	16.0 (14.3, 17.3)	0.91 (0.86, 0.95)	-1.49 (-2.31, -0.86)
(23) 9 + 17 + 18	4.9 (3.3, 6.8)	0.62 (0.43, 0.82)	-3.00 (-4.44, -1.39)	11.0 (8.5, 13.8)	0.63 (0.50, 0.78)	-6.47 (-8.51, -3.92)
(24) 10 + 14 + 18	7.2 (6.2, 8.1)	0.91 (0.84, 0.95)	-0.72 (-1.30, -0.39)	15.8 (14.3, 17.5)	0.91 (0.87, 0.95)	-1.63 (-2.31, -0.84)
(25) 13 + 17 + 18	4.0 (2.5, 5.9)	0.51 (0.33, 0.72)	-3.91 (-5.26, -2.18)	11.0 (8.3, 14.1)	0.63 (0.49, 0.80)	-6.44 (-8.72, -3.50)

Author Manuscript

Author Manuscript

Author Manuscript

Author Manuscript

Abbreviations: hPDI (healthy plant-based dietary index score), ACS (American Cancer Society nutrition guidelines score), HEI (Healthy Eating Index score), PA (moderate to vigorous physical activity in hrs/wk), RR (risk ratio), RD (risk difference), CI (confidence interval).

^a Adjusted for age at diagnosis, race/ethnicity, education, body mass index, alcohol intake, energy intake, comorbidity score, cancer stage, estrogen receptor status, progesterone receptor status, human epidermal growth factor receptor 2, surgery type, chemo, radiation and hormone therapies (see footnotes ^{d,f,g} for additional covariates included depending on the intervention).

^b Reported per 100 participants.

^c Diet quality cut points determined from the 25th, 50th, 75th and 100th percentiles of the baseline distributions for each dietary index.

^d In addition to covariates listed in footnote a, physical activity and smoking were added to the models as covariates.

^e Physical activity cut points determined from Physical Activity Guidelines for Americans as set forth by the United States Department of Health and Human Services (see text for more details).

^f In addition to covariates listed in footnote a, HEI and smoking were added to the model as covariates.

^g In addition to covariates listed in footnote a, HEI and physical activity were added to the model as covariates.