The influence of a breast cancer diagnosis on serum 25-hydroxyvitamin D

Katie M. O'Brien, Dale P. Sandler, Melissa House, Jack A. Taylor, and Clarice R. Weinberg

Correspondence to Dr. Katie O'Brien, Epidemiology Branch, National Institute of Environmental Health Sciences,111 TW Alexander Dr., Research Triangle Park, NC 27709 (email: obrienkm2@niehs.nih.gov) phone: 984-287-3753

Author Affiliations: Epidemiology Branch, National Institutes of Environmental Health Sciences (Katie M. O'Brien, Dale P. Sandler, Jack A. Taylor); Biostatistics and Computational Biology Branch, National Institute of Environmental Health Sciences (Katie M. O'Brien, Clarice R. Weinberg); Westat, Inc., Durham, NC (Melissa House).

Funding: This work was supported by an Office of Dietary Supplement Research Scholars Program Grant (to KMO) and the Intramural Research Program of the National Institutes of Health, National Institute of Environmental Health Sciences (project Z01-ES044005 to DPS and Z01-ES102245 to CRW).

Conflict of Interests: None declared.

Running Title: Vitamin D Over Time and Breast Cancer Risk

Abbreviations: 25(OH)D = 25-hydroxyvitamin D; BMI = Body Mass Index; CI= Confidence Interval; ICC = intraclass correlation coefficients; OR= Odds Ratio

Abstract

Prospective and retrospective studies of vitamin D and breast cancer have produced discrepant results. This may be due to variations in serum 25-hydroxyvitamin D [25(OH)D] concentrations over time, including systematic changes after breast cancer diagnosis. We measured total serum 25(OH)D in Sister Study participants who provided samples at baseline (2003-2009) and 4-10 years later (2013-2015). This included 827 women with an intervening breast cancer and 771 without. Although modestly correlated over time (R=0.42), 25(OH)D concentrations increased in both groups, with larger increases among cases (averaging 31.6 ng/mL at baseline, 43.5 ng/mL at follow-up) than controls (32.3 ng/mL at baseline, 40.4 ng/mL at follow-up). Consequently, the estimated association between 25(OH)D and breast cancer depended on whether baseline (odds ratio [OR]=0.87, 95% confidence interval [CI]: 0.78-0.98 per 10 ng/mL) or second blood draw measures (OR=1.17, 95% CI: 1.08-1.26 per 10 ng/mL) were used. Concentrations were related to regular (≥4 times/week) vitamin D supplement use, which became more common over time. Increases were greater in cases (56% to 84%) than in controls (56% to 77%). Our results do not explain previously observed differences between retrospective and prospective studies, but do demonstrate how reverse causation and temporal trends in exposure can distort inference.

Keywords: vitamin D; breast cancer; 25-hydroxyvitamin D; reverse causation bias; reliability

Introduction

Vitamin D has known, anti-carcinogenic properties (1), but previous clinical trials and observational studies have not established a clear association between vitamin D and breast cancer risk (2–4). Observational studies of the association between breast cancer and the vitamin D biomarker, 25-hydroxyvitamin D [25(OH)D], have produced mixed results. Although many retrospective case-control studies that measured 25(OH)D after the cases had been diagnosed reported evidence suggesting that 25(OH)D is inversely associated with breast cancer (5–9), results from prospective cohort studies have been less consistent. In most, pre-diagnostic 25(OH)D concentrations were either not associated with or only weakly inversely associated with breast cancer (2,10,19–25,11–18).

Our recent prospective study of serum 25(OH)D and breast cancer risk (26) found that high concentrations of 25(OH)D were associated with decreased risk of breast cancer over the subsequent five years. We had restricted the follow up interval because 25(OH)D levels vary over time (27–29) and recent exposure could be most relevant to risk, with attenuation of effect estimates over prolonged follow-up. We further hypothesized that the protective association observed in retrospective case-control studies could also be due, in part, to reverse causation, with recently diagnosed cases experiencing lifestyle changes or treatment effects that reduced their 25(OH)D concentrations. However, as we only had measures at a single time point prior to cases' diagnoses, we could not estimate effects of either time or of breast cancer on 25(OH)D.

We undertook this new study to address those remaining gaps. Specifically, we aimed to assess serum 25(OH)D in samples collected several years after baseline to study changes over time. If changes did occur, we wanted to understand what factors influenced these changes, with particular attention to an intervening breast cancer diagnosis.

Methods

Description of Study Participants

Participants were sampled from the prospective Sister Study cohort, which included 50,884 women who had never had breast cancer when they enrolled. To be eligible, women had to have at least one sister previously diagnosed with breast cancer, be aged 35-74, and live in the United States. During enrollment (2003-2009), participants completed a computer-assisted telephone interview, providing data on demographics, reproductive history, personal and familial health histories, and other topics. Trained medical examiners collected bio-specimens and took body measurements and all participants provided written informed consent. The study was approved and overseen by the Institutional Review Boards of the National Institute of Environmental Health Sciences and the Copernicus Group.

Participants are re-contacted annually for updates on their health. More detailed questionnaires are completed every 2-3 years. Women with incident cancer are asked to provide a copy of the pathology report and authorize release of their medical records. All analyses are based on data collected through September 2016 (data release 6.0).

We previously assessed baseline 25(OH)D concentrations in 3,386 women for a case-cohort study of incident breast cancer (26). This included 1,843 women randomly selected from the cohort (including 68 cases) and an additional 1,543 women who had been diagnosed with ductal carcinoma *in situ* or invasive breast cancer within five years of baseline. During 2013-2015, 3,707 women who had provided baseline blood samples (1,838 who had developed breast cancer and 1,869 women who had not) were asked to provide a second blood sample. We successfully collected second samples for 1,144 cases (62%) and 1,214 controls (65%). Of these, 1598 (827 cases, 771 controls) with baseline 25(OH)D measures were randomly selected for inclusion in this study of 25(OH)D over time.

Assessment of serum 25(OH)D

The same protocol was used to collect, store, ship and analyze baseline and follow-up serum samples for 25(OH)D (26). Briefly, blood was collected during in-home exams and shipped overnight to our laboratory for aliquoting (30). All samples were stored at -80˚ C in 0.4-0.5 mL straws before being shipped to Heartland Assays (Ames, IA). Once there, samples were analyzed for $25(OH)D_3$, $25(OH)D_2$ and 3-epi-$25(OH)D_3$ using liquid chromatography–mass spectrometry with an Agilent 1290 Series High-Pressure Liquid Chromatography system and an Agilent 6460 Triple Quadruple liquid chromatography–mass spectrometry. We summed the concentrations of the three metabolites, using that total as our measure of interest. Individual metabolite concentrations below the limit of detection (1.5 ng/mL) were assigned a value of 1.06 ng/mL (=1.5/√2). This occurred for 0%, 69% and 26% of $25(OH)D_3$, $25(OH)D_2$ and 3-epi-$25(OH)D_3$, respectively, at baseline, and 0%, 85% and 21% of $25(OH)D_3$, $25(OH)D_2$ and 3-epi-$25(OH)D_3$ at follow-up. We adjusted for batch and season of blood draw using random effects models and LOESS regression, respectively, with the LOESS analyses stratified by race (African-American versus non-African-American) and regular vitamin D supplement use (≥4 days/week versus <4 days/week).

The baseline and second blood samples were analyzed separately, in 2015 and 2017, respectively. For quality control across batches and over time, each batch included two samples drawn from pooled sera. One of the pools combined sera from premenopausal women and the other combined sera from postmenopausal women. For the 2015 analysis, we observed inter-batch coefficients of variation 11.0% and 8.5% for the premenopausal and postmenopausal samples, respectively. Precision improved for the 2017 samples, with coefficients of variation of 5.4% and 4.9%. The QC samples showed slightly lower values for the 2017 versus 2015 assays (Fisher's combined p-value for t-tests p=0.07).

Statistical Analysis

Correlation over time

We first calculated Spearman's rank correlation coefficients (R) and p-values to compare within-individual total 25(OH)D concentrations at the two time points. This was done for all samples together and then separately for cases and controls. We also calculated intraclass correlation coefficients (ICCs).

Predictors of changes over time

We next constructed predictive models to assess how within-individual changes in certain factors over time were related to within-individuals changes in 25(OH)D concentrations. The covariates considered included an intervening breast cancer diagnosis, time between sample collections, average age across the study period, and changes in the following variables across the two times of blood collection: hormonal birth control use, hormone therapy use, physical activity level, body mass index (BMI), alcohol use, history of osteoporosis, menopausal status, vitamin D supplement use, time spent outdoors, smoking status, geographic location, waist circumference, waist to hip ratio, use of protective clothing, and incident bone fractures. We did not adjust for baseline 25(OH)D, which would have produced bias towards the mean. For case-only models, we additionally considered disease-related factors, including stage (0/*in situ*, I, II, III/IV), type (lobular or ductal), estrogen receptor status, progesterone receptor status, and human epidermal growth factor receptor status, as well as treatment type: chemotherapy (yes/no), Herceptin/biological agents (yes/no), radiation (yes/no), hormonal treatment (yes/no), and type of surgery (none, mastectomy, lumpectomy or other breast conserving surgery). These were entered into a linear regression model in a stepwise fashion (with p-value cut-points of 0.15 to enter and 0.10 to stay).

We also constructed predictive models to better assess 25(OH)D changes over calendar time versus age time. These models were fit separately for cases and controls and used restricted cubic

spline terms for age and year. We used generalized estimating equations to account for repeated measures (two per individual). To assess age trends, we obtained the predicted 25(OH)D concentration for specified age values (range 35-85) when year was as observed. This was repeated to look at the predicted 25(OH)D concentration for each year (2003-2015), when age was as observed.

Modeling 25(OH)D and breast cancer risk

We next used logistic regression to assess the relationship between different measures of 25(OH)D and breast cancer risk. For the first set of models, we estimated odds ratios (ORs) and 95% confidence intervals (CIs) for the association between breast cancer and baseline 25(OH)D, both continuously and categorized into quartiles (820 cases, 764 controls with complete covariate data). We also estimated ORs for the relationship between regular vitamin D supplement use (a vitamin D-containing supplement ≥4 times/week) and breast cancer risk. These models were adjusted for the following covariates, as measured at baseline: age at blood draw (continuous), race/ethnicity (non-Hispanic white, Black, Hispanic or other), education (≤high school, some college, bachelor's degree, graduate degree), current hormonal birth control use (yes/no), current hormone therapy use (none, estrogen plus progestin, estrogen alone), physical activity level (0-5 hours/week, >5-10 hours/week, >10 hours/week), BMI (<25 kg/m^2, 25-29.9 kg/m^2, ≥30kg/m^2), alcohol use (never/former drinker, current drinker <1 drink/day, current drinker ≥1 drink/day), history of osteoporosis (yes/no), parity (0, 1, 2, ≥3 births), menopausal status (pre or post-menopausal), and a BMI by menopausal status interaction term.

In a second set of logistic regression analyses, we used vitamin D measures based instead on the second blood draw, but adjusted for the confounders at their baseline levels. This was meant to represent a retrospective case-control study where exposure is assessed after cases were diagnosed, but participants are asked to recall covariate information from a prior time point. Lastly, we calculated

ORs and 95% CIs for the association between vitamin D measures at the second blood draw, adjusting for covariates defined as of the time of that visit (as determined by data from follow-up questionnaires; including 822 cases, 765 controls with complete covariate data). As very few women reported current hormonal birth control use at the time of second blood draw, this was omitted from the model. We additionally combined all types of hormone therapy into an ever/never variable.

Sensitivity Analyses

In the first of two sensitivity analyses, we stratified cases by time between breast cancer diagnosis and second blood draw (<1, 1-2.9, ≥3 years). Each set of cases was compared to all controls using logistic regression. In a second sensitivity analysis, we addressed possible selection bias by weighting the included participants by their inverse probability of selection (31,32), relative to the subset asked to participate in the nested case-control biospecimen study. The weights for these analyses were calculated separately for cases and controls using logistic regression and included baseline levels of the variables specified for the stepwise regression plus disease-related factors for the cases.

Results

The most notable difference between cases and controls (Table 1) was that although concentrations increased substantially over time in both groups, controls had higher relative 25(OH)D concentrations at baseline (mean=32.3 ng/mL versus 31.6 in cases; Figure 1A), but lower concentrations than cases at the second blood draw (40.4 ng/mL in controls versus 43.5 ng/mL in cases; Figure 1B; $p=7x10^{-9}$ for the estimated effect of case status on changes over time). The increase over time is consistent with the increasing proportion of women reporting use of a vitamin D-

containing supplement at least four times per week (56% at baseline for both groups, but 77% and 84% for controls and cases, respectively, at time of second blood draw).

Waist circumferences increased over time for both groups, with increasing BMI seen in controls, an increasing proportion post-menopausal (especially among cases), and a slight decrease in current smokers in both groups. Other variables, including physical activity, time spent outdoors, and alcohol consumption were difficult to compare across time due to differences in how the questions were assessed at baseline versus follow-up (see Web Appendix, Web Figures 1-6), but are still useful for comparing cases versus controls.

The cases were diagnosed at age 60.3, on average (Web Table 1). Most were estrogen receptor positive, progesterone receptor positive, and human epidermal growth factor receptor-2 negative and 75% were either stage 0 (*in situ*) or stage I. Eight-two percent had either a mastectomy or lumpectomy, 72% were treated with a hormonal agent such as Tamoxifen or Arimidex, 65% received radiation and 36% received chemotherapy.

Correlation over time

Total 25(OH)D concentrations were modestly correlated over time ($R=0.42$, $p<0.001$; Web Figure 7). ICCs were 0.18, 0.09 and 0.29 for everyone, cases, and controls, respectively.

Predictors of changes over time

Table 2 shows the final prediction model for changes in 25(OH)D over time, with each estimate adjusted for the other covariates listed. Breast cancer status was one of the main determinants of changes in 25(OH)D over time among all participants with an estimated 3.26 ng/mL (95% CI: 1.96, 4.56) increase associated with an intervening diagnosis. Other key contributors were a longer time between blood draws, discontinuing hormonal birth control use, initiating or discontinuing regular vitamin D supplement use, older average age between blood draws, and increasing alcohol

consumption. Among controls, time between blood draws, discontinuing hormonal birth control use, average age and vitamin D supplementation were again important predictors. Among cases, time between blood draws, discontinuing hormonal birth controls use, regular vitamin D supplement use, and alcohol consumption predicted 25(OH)D. None of the disease or treatment related factors were retained in the stepwise case-only model. Estimated parameters for the initial models with all considered covariates are shown in Web Tables 2-4.

When we modeled the predicted effects of age and calendar time, we observed that both were positively associated with 25(OH)D levels in cases and controls after mutual adjustment (Web Figure 8). In both instances, the slope for calendar time was steeper than for age, indicating the observed changes were primarily behavioral.

Modeling 25(OH)D and breast cancer risk

Similar to our previous report (26), total baseline 25(OH)D concentrations were inversely associated with breast cancer risk (OR=0.87, CI: 0.78-0.98 per 10 ng/mL increase; Table 3), with some evidence of a threshold effect for levels in the fourth quartile (OR=0.71, 95% CI: 0.52, 0.98 for 25(OH)D >38.5 [4th quartile] versus ≤25.3 ng/mL [1st quartile]). We also observed an inverse association between regular vitamin D supplement use and breast cancer risk (OR=0.87, 95% CI: 0.70, 1.07). The associations between vitamin D and breast cancer were similar in analyses restricted to women who were post-menopausal at baseline (Web Table 5).

In contrast to these findings, when we examined the relationship between breast cancer status and vitamin D measured at the second time point, adjusting for contemporary covariates, we observed strong positive associations. This was true for the analysis of total 25(OH)D measured continuously (OR=1.17, 95% CI: 1.08, 1.26) or categorized into quartiles, as well as the assessment of regular supplement use. Results were nearly identical for models adjusted for covariates measured at baseline.

Sensitivity Analyses

The association between 25(OH)D at the second blood draw and breast cancer did not measurably depend on the time between diagnosis and second blood draw (Web Table 6). The positive association between regular supplement use at the time of the second blood draw and breast cancer was also consistent over time. Of the women who were invited to participate in the nested case-control bio-specimen study, those who participated tended to be older, non-Hispanic white, better educated, have lower BMI, and were less likely to be smokers (Web Table 7). When we used inverse probability of selection weights to control this selection bias, there was little difference in the effect estimates for 25(OH)D (Web Table 8) and a slightly stronger effect estimate for regular supplement use (HR=1.54, 95% CI: 1.18, 2.01).

Discussion

In this study of serum 25(OH)D concentrations over time and their relationship with breast cancer, we found that 25(OH)D was modestly correlated over a 4-10 year period. However, concentrations increased considerably between baseline (2003-2009) and second blood draws (2013-2014), with a larger increase seen in cases than controls. These increases mirrored increases in participants' self-reported regular vitamin D supplement use, again with a higher increase seen in those with an intervening diagnosis of breast cancer.

In accordance with these trends, although baseline 25(OH)D was associated with a decreased risk of breast cancer, the later 25(OH)D levels were associated with increased risk. These findings do not help to explain the previously observed differences between retrospective and prospective studies of 25(OH)D and breast cancer, in which only retrospective studies tended to show protection.

However, they do demonstrate that reverse causation and temporal trends in exposure can hugely bias effect estimates in retrospective studies.

Our findings that vitamin D supplement use and 25(OH)D concentrations increased over the study period are consistent with data collected for the National Health and Nutrition Survey (NHANES). For 1999-2012, Kantor et al. (33) observed a substantial increase in vitamin D supplementation in women, going from 40% prevalence in 1999-2000 to 47% prevalence in 2011-2012. Much of the change seemed to be driven by an increase in vitamin D-specific supplementation (8% to 26% prevalence over the 14-year period) rather than from general multivitamin intake, which actually decreased (44% to 41%). As the initial prevalence and slope of change in prevalence over time were highest for older, well-educated, and non-Hispanic white women, these nationally-representative results are quite consistent with what was observed in our sample.

Though consistently lower than in Sister Study participants, 25(OH)D concentrations in NHANES participants also rose over this time period, from a mean of 24.4 ng/mL in 2003-2004 (unadjusted for season and limited to females >35) to a mean of 31.0 ng/mL in 2011-2012. (34) Based on our data, the increase in vitamin D use may be stronger in women diagnosed with breast cancer. Such women are presumably especially motivated to improve their overall health and survival, decrease their risk of recurrence, and manage treatment side effects.

Our correlation results are lower than other studies of 25(OH)D concentrations over time. Not surprisingly, the studies with shorter time intervals between samples generally had higher correlations than those with longer time periods. Samples collected one year apart had correlations of 0.65-0.80 (28,35) or ICCs of 0.90 (36); 1-3 years apart had ICCs of 0.68-0.96; (24,36,37) longer-term studies showed ICCs of 0.50-0.65 (24,27) or correlation coefficients of 0.42-0.53 (28,35). However, those studies preceded the recent rise in vitamin D supplement use.

Discontinuation of hormonal birth control was associated with decreased 25(OH)D. However, despite the fact that women were aging and going through menopause (including treatment-induced menopause) over the study period, 25(OH)D still increased substantially. Given that neither dietary intake nor time spent outdoors were retained in the stepwise prediction models, it seems likely that increased supplement use is the principal driver of 25(OH)D increases in our participants.

Our initial expectation was that women recently diagnosed with breast cancer would have lower 25(OH)D concentrations due to the effects of the disease, its treatment, or to behavioral changes following diagnosis. Thus, we were surprised to see positive associations between retrospectively-collected 25(OH)D and breast cancer – even among those who provided their second serum sample within one year of diagnosis. It is possible that the second samples were collected too long after diagnosis to capture transient disease-related decreases. After all, the previous case-control studies that showed inverse associations between retrospectively-measured 25(OH)D and breast cancer typically enrolled participants within a few months of their diagnosis. (5–7,9) As such, they may have done a better job of capturing recent vitamin D exposure, in which case the previous results may reflect our hypothesis that recent vitamin D exposure is most relevant to breast carcinogenesis. That said, we note that these previous case-control studies also predate the recent increases in vitamin D supplement use (33). Case-control studies may also have suffered from selection bias, with participating controls more health conscious and thus more likely to take supplements than the population from which they were drawn.

A limitation of this study is that the assays for the baseline and follow-up 25(OH)D measures were done at two separate times, thereby increasing overall variability. However, we used the same company and analysis approach for both sets of samples and included some of the same pooled control samples with both sets of assays. Though comparisons of the quality control samples did not show

statistically-significant differences, there was some weak evidence of systematic differences. If there is a true systematic bias between the two sets of tests, the differences are likely small and should be non-differential by case status.

Generalizability is also a concern here, as our participants are mostly non-Hispanic white and well-educated, and all have a sister with a history of breast cancer. Such women likely have high vitamin D intakes than the general population, though their reported supplement use was consistent with the trends seen in NHANES participants of similar age, race, and education. Additionally, as the response rates for the second blood draws were somewhat low (63%), we examined estimated effects for the original selected sample using inverse probability of selection weights. These analyses produced results similar to those reported in the main analysis.

Study strengths include the use of liquid chromatography–mass spectrometry to measure 25(OH)D and detailed and frequent data collection that allowed us to assess exposure and covariate levels across the study period. This is also the first study to compare trends for women with and without intervening breast cancer diagnoses. With this data we were able to compare results obtained from looking at the 25(OH)D-breast cancer association prospectively (pre-diagnosis) versus retrospectively (post-diagnosis). Though our finding of an apparent positive association between 25(OH)D and breast cancer in the retrospective analysis using samples taken at follow-up was not what we had expected, the increased use of vitamin D supplements over the study period, especially among cases, offers a plausible explanation for the observed effect estimates in this specific case. More generally, this study provides a cautionary tale supporting the need to consider temporal trends in exposure, especially when those exposures may be affected by disease status.

Figure 1. Distribution of serum 25(OH)D concentrations by breast cancer status at A) baseline and B) time of second blood draw. At baseline, the mean 25(OH)D concentration for cases was 31.6 ng/mL and the mean for controls was 32.3 ng/mL. At follow-up, the means were 43.5 ng/mL and 40.4 ng/mL for cases and controls, respectively.

Author Affiliations: Epidemiology Branch, National Institutes of Environmental Health Sciences (Katie M. O'Brien, Dale P. Sandler, Jack A. Taylor); Biostatistics and Computational Biology Branch, National Institute of Environmental Health Sciences (Katie M. O'Brien, Clarice R. Weinberg); Westat, Inc., Durham, NC (Melissa House).

Funding: This work was supported by an Office of Dietary Supplement Research Scholars Program Grant (to KMO) and the Intramural Research Program of the National Institutes of Health, National Institute of Environmental Health Sciences (project Z01-ES044005 to DPS and Z01-ES102245 to CRW).

Acknowledgements: The authors would like to thank Cindy Kleeberger for her assistance with managing the biospecimens and Drs. Srishti Shrestha and Kaitlyn Gam for their review of an early version of this manuscript.

Conflict of Interests: The authors declare they have no actual or potential competing financial interests.

References

1. Holick MF. High prevalence of vitamin D inadequacy and implications for health. *Mayo Clin. Proc.* 2006;81(3):353–373.
2. Chlebowski RT, Johnson KC, Kooperberg C, et al. Calcium plus vitamin D supplementation and the risk of breast cancer. *J Natl Cancer Inst*. 2008;100(22):1581–1591.
3. Gandini S, Boniol M, Haukka J, et al. Meta-analysis of observational studies of serum 25-hydroxyvitamin D levels and colorectal, breast and prostate cancer and colorectal adenoma. *Int. J. Cancer*. 2011;128(6):1414–1424.
4. Bolland MJ, Grey A, Gamble GD, et al. Calcium and vitamin D supplements and health outcomes: a reanalysis of the Women's Health Initiative (WHI) limited-access data set. *Am J Clin Nutr*. 2011;94(4):1144–1149.
5. Abbas S, Linseisen J, Slanger T, et al. Serum 25-hydroxyvitamin D and risk of post-menopausal breast cancer--results of a large case-control study. *Carcinogenesis*. 2008;29(1):93–99.
6. Abbas S, Chang-Claude J, Linseisen J. Plasma 25-hydroxyvitamin D and premenopausal breast cancer risk in a German case-control study. *Int J Cancer*. 2009;124(1):250–255.
7. Chen P, Li M, Gu X, et al. Higher blood 25(OH)D level may reduce the breast cancer risk: evidence from a Chinese population based case-control study and meta-analysis of the

observational studies. *PLoS One*. 2013;8(1):e49312.

8. Colston KW, Lowe LC, Mansi JL, et al. Vitamin D status and breast cancer risk. *Anticancer Res*. 2006;26(4A):2573–2580.
9. Crew KD, Gammon MD, Steck SE, et al. Association between plasma 25-hydroxyvitamin D and breast cancer risk. *Cancer Prev Res*. 2009;2(6):598–604.
10. Almquist M, Bondeson A-G, Bondeson L, et al. Serum levels of vitamin D, PTH and calcium and breast cancer risk-a prospective nested case-control study. *Int J Cancer*. 2010;127(9):2159–2168.
11. Amir E, Cecchini RS, Ganz PA, et al. 25-Hydroxy vitamin-D, obesity, and associated variables as predictors of breast cancer risk and tamoxifen benefit in NSABP-P1. *Breast Cancer Res Treat*. 2012;133(3):1077–1088.
12. Bertone-Johnson ER, Chen WY, Holick MF, et al. Plasma 25-hydroxyvitamin D and 1,25-dihydroxyvitamin D and risk of breast cancer. *Cancer Epidemiol Biomarkers Prev*. 2005;14(8):1991–1997.
13. Deschasaux M, Souberbielle J, Latino-martel P, et al. Weight Status and Alcohol Intake Modify the Association between Vitamin D and Breast Cancer Risk. *J Nutr*. 2016;(C):1–10.
14. Eliassen AH, Spiegelman D, Hollis BW, et al. Plasma 25-hydroxyvitamin D and risk of breast cancer in the Nurses' Health Study II. *Breast Cancer Res*. 2011;13(3):R50.
15. Eliassen AH, Warner ET, Rosner B, et al. Plasma 25-hydroxyvitamin D and risk of breast cancer in women followed over 20 years. *Cancer Res*. 2016;76(18):5423–5430.
16. Engel P, Fagherazzi G, Boutten A, et al. Serum 25(OH) vitamin D and risk of breast cancer: A nested case-control study from the French E3N cohort. *Cancer Epidemiol. Biomarkers Prev.* 2010;19(9):2341–2350.
17. Freedman DM, Chang S-CS-C, Falk RT, et al. Serum levels of vitamin D metabolites and breast cancer risk in the prostate, lung, colorectal, and ovarian cancer screening trial. *Cancer Epidemiol. Biomarkers Prev. ()(pp 889-894), 2008.Date Publ. April 2008.* 2008;17(United States RF-33 LG-English PT-Journal: Article DD-20080514):889–894.
18. Kim Y, Franke AA, Shvetsov YB, et al. Plasma 25-hydroxyvitamin D3 is associated with decreased risk of postmenopausal breast cancer in whites: a nested case-control study in the multiethnic cohort study. *BMC Cancer*. 2014;14:29.
19. Kuhn T, Kaaks R, Becker S, et al. Plasma 25-hydroxyvitamin D and the risk of breast cancer in

the European prospective investigation into cancer and nutrition: a nested case-control study. *Int J Cancer*. 2013;133(7):1689–1700.

20. McCullough ML, Stevens VL, Patel R, et al. Serum 25-hydroxyvitamin D concentrations and postmenopausal breast cancer risk: a nested case control study in the Cancer Prevention Study-II Nutrition Cohort. *Breast Cancer Res* [electronic article]. 2009;11(4):R64. (http://breast-cancer-research.biomedcentral.com/articles/10.1186/bcr2356)
21. Mohr SB, Gorham ED, Alcaraz JE, et al. Serum 25-hydroxyvitamin D and breast cancer in the military: a case-control study utilizing pre-diagnostic serum. *Cancer Causes Control*. 2013;24(3):495–504.
22. Ordóñez-Mena JM, Schöttker B, Fedirko V, et al. Pre-diagnostic vitamin D concentrations and cancer risks in older individuals: an analysis of cohorts participating in the CHANCES consortium. *Eur J Epidemiol*. 2016;31:311–323.
23. Rejnmark L, Tietze A, Vestergaard P, et al. Reduced prediagnostic 25-hydroxyvitamin D levels in women with breast cancer: a nested case-control study. *Cancer Epidemiol Biomarkers Prev*. 2009;18(10):2655–2660.
24. Scarmo S, Afanasyeva Y, Lenner P, et al. Circulating levels of 25-hydroxyvitamin D and risk of breast cancer: a nested case-control study. *Breast Cancer Res*. 2013;15:R15.
25. Skaaby T, Husemoen LLN, Thuesen BHH, et al. Prospective Population-Based Study of the Association between Serum 25-Hydroxyvitamin-D Levels and the Incidence of Specific Types of Cancer. *Cancer Epidemiol Biomarkers Prev*. 2014;23(7):1220–1229.
26. O'Brien KM, Sandler DP, Taylor JA, et al. Serum Vitamin D and Risk of Breast Cancer within Five Years. *Environ. Health Perspect.* 2017;125(7):077004.
27. Bertrand KA, Giovannucci E, Liu Y, et al. Determinants of plasma 25-hydroxyvitamin D and development of prediction models in three US cohorts. *Br J Nutr*. 2012;108(10):1889–1896.
28. Hofmann JN, Yu K, Horst RL, et al. Long-term variation in serum 25-hydroxyvitamin D concentration among participants in the Prostate, Lung, Colorectal, and Ovarian Cancer Screening Trial. *Cancer Epidemiol. Biomarkers Prev.* 2010;19(4):927–31.
29. Knight JA, Wong J, Cole DEC, et al. Predictors of 25-hydroxyvitamin D measured at multiple time points in a multiethnic population. *Am J Epidemiol*. 2017;186(10):1180–1193.
30. Sandler DP, Hodgson ME, Deming-Halverson SL, et al. The Sister Study: Baseline methods and participant characteristics. *Env. Heal. Perspect*. 2017;125(12):127003.

31. Cole SR, Hernán MA. Constructing inverse probability weights for marginal structural models. *Am. J. Epidemiol.* 2008;168(6):656–664.
32. Weisskopf MG, Sparrow D, Hu H, et al. Biased Exposure-Health Effect Estimates from Selection in Cohort Studies: Are Environmental Studies at Particular Risk? *Env. Heal. Perspect.* 2015;123(11):1113–1122.
33. Kantor ED, Rehm CD, Du M, et al. Trends in Dietary Supplement Use Among US Adults From 1999-2012. *Jama* [electronic article]. 2016;316(14):1464. (http://dx.doi.org/10.1001/jama.2016.14403)
34. Center For Disease Control. National Health and Nutrition Examination Survey, National Center for Health Statistics. (https://www.cdc.gov/nchs/nhanes/index.htm). (Accessed January 26, 2018)
35. Jorde R, Sneve M, Hutchinson M, et al. Tracking of serum 25-hydroxyvitamin D levels during 14 years in a population-based study and during 12 months in an intervention study. *Am J Epidemiol*. 2010;171(8):903–908.
36. Sonderman JS, Munro HM, Blot WJ, et al. Reproducibility of serum 25-Hydroxyvitamin D and vitamin D-binding protein levels over time in a prospective cohort study of black and white adults. *Am. J. Epidemiol.* 2012;176(7):615–621.
37. Kotsopoulos J, Tworoger SS, Campos H, et al. Reproducibility of plasma and urine biomarkers among premenopausal and postmenopausal women from the nurses health studies. *Cancer Epidemiol. Biomarkers Prev.* 2010;19(4):938–946.

Table 1. Characteristics of Sister Study participants at baseline (2003-2009) and second blood draw (2013-2015).[a]

Characteristic	Controls (n=771)				Breast cancer cases (n=827)			
	Baseline		Second blood draw		Baseline		Second blood draw	
	No.	%	No.	%	No.	%	No.	%
Age[b]	55.6 (8.8)		63.1 (8.8)		57.8 (8.8)		65.4 (8.7)	
Total 25(OH)D (ng/mL)[b]	32.3 (10.3)		40.4 (13.7)		31.6 (9.7)		43.5 (14.1)	
Waist in inches[b]	33.7 (5.7)		35.1 (5.9)		34.1 (5.8)		35.3 (5.8)	
Waist to hip ratio[b]	0.8 (0.1)		0.8 (0.1)		0.8 (0.1)		0.8 (0.1)	
Race								
Non-Hispanic white	685	89			737	89		
African-American	46	6			38	5		
Hispanic	27	4			27	3		
Other	13	2			24	3		
Education								
High school or less	110	14			118	14		
Some college	241	31			237	29		
Bachelor's degree	203	26			214	26		
Graduate degree	217	28			257	31		
Postmenopausal	513	67	678	88	593	72	788	95
Physical activity[c]								
0-5 hours/week	233	30	332	43	233	28	337	41
5.1 – 10 hours/week	279	36	187	24	314	38	222	27
>10 hours/week	258	34	252	33	278	34	268	32
Body mass index[d]								
<25 kg/m^2	305	40	292	38	321	39	303	37
25-29.9 kg/m^2	247	32	237	31	265	32	283	34
≥30 kg/m^2	219	28	242	31	241	29	241	29
Current hormonal birth control	34	4	12	2	33	4	0	0
Current hormone therapy								
None	682	89	607	79	706	86	723	88
Estrogen plus progestin	26	3	32	4	47	6	15	2
Estrogen only	62	8	127	17	72	9	85	10
Regular vitamin D supplement use (≥4x/week)	425	56	590	77	456	56	691	84
Time Spent Outdoors[c]								
0-320 hours/year	198	26	179	23	209	25	230	28
321-530 hours/year	195	25	151	20	198	24	139	17
531-850 hours/year	184	24	262	34	215	26	301	36
>850 hours/year	193	25	179	23	203	25	157	19
Usually wear protective clothing/hat[a]	169	22	254	33	181	22	276	33
Alcohol Consumption[c]								
Never/former drinker	131	17	162	21	150	18	192	23
Current drinker, <1 drink/day	519	67	430	56	566	68	486	59
Current drinker, ≥1 drink/day	119	15	179	23	111	13	149	18
Smoking Status								
Never Smoker	417	54	417	54	466	56	466	56
Former Smoker	306	40	324	42	324	39	337	41
Current Smoker	48	6	30	4	37	4	24	3
Parity								
0 births	154	20	151	20	171	21	171	21
1 birth	110	14	112	15	110	13	109	13
2 births	277	36	278	36	291	35	292	35

≥3 births	229	30	229	30	255	31	255	31
History of osteoporosis	190	25	218	28	190	23	245	30
State of residence								
1-Southern most (FL, HI, PR)	54	7	57	7	62	7	63	8
2	104	13	107	14	106	13	111	13
3	250	32	255	33	264	32	260	31
4	288	37	278	36	315	38	314	38
5- Northern most (AK, WA, MT, ND, MN, WI, ME)	75	10	74	10	80	10	79	10

Abbreviations: FL = Florida, HI= Hawaii, PR = Puerto Rico, AK = Alaska, WA = Washington, MT = Montana, ND = North Dakota, MN = Minnesota, WI= Wisconsin, and ME = Maine

[a]Whenever possible, missing data at follow-up filled in based on baseline data (<1% of observations). Remaining missing: Baseline waist (2 cases), Baseline waist to hip ratio (2 controls, 2 cases), Race/ethnicity (1 case), Education (1 case), Physical activity at baseline (1 control, 2 cases), Current birth control use at baseline (2 controls, 2 cases), Current hormone therapy use at baseline (1 control, 2 cases), Current hormone therapy at follow-up (5 controls, 4 cases), Regular vitamin D supplement use at baseline (13 controls, 9 cases), Regular vitamin D supplement use at follow-up (5 controls, 4 cases), Time spent outdoors at baseline (1 control, 2 cases), Alcohol consumption at baseline (2 controls), Parity at baseline (1 control), Parity at follow-up (1 control)

[b]Values are expressed as mean (standard deviation)

[c]Variables assessed differently in baseline and follow-up questionnaires. See web appendix for details.

[d]Measured as weight (kg) / height (m)^2

Table 2. Parameters retained from stepwise regression models for predicting changes in total serum 25(OH)D levels over time.

Parameter	Change in 25(OH)D per unit increase					
	All Participants		Controls		Cases	
	β	95% CI	β	95% CI	β	95% CI
Breast Cancer Event	3.26	1.96, 4.56				
Time between blood draws (years)	1.01	0.53, 1.49	0.98	0.33, 1.64	0.95	0.26, 1.64
Stop taking hormonal birth control	-6.16	-9.62, -2.69	-5.83	-10.7, -0.94	-7.58	-12.3, -2.86
Starting taking a vitamin D supplement regularly (≥4 times/week)	3.37	1.92, 4.82	4.48	2.41, 6.55	2.35	0.31, 4.38
Stopped taking a vitamin D supplement regularly	-8.30	-11.8, -4.79	-7.24	-11.7, -2.76	-9.94	-15.5, -4.41
Average age across blood draws (years)	0.08	0.01, 0.16	0.10	0.00, 0.21		
Increase in alcohol consumption (per category)	-1.49	-2.76, -0.21			-2.31	-4.14, -0.49

CI = confidence interval

Table 3. Odds ratios and 95% confidence intervals for relationship between vitamin D and breast cancer based on different measurements of vitamin D and associated covariates

Vitamin D category	Odds Ratio	95% Confidence Interval
Model 1: Baseline 25(OH)D D, Baseline covariates (820 cases, 764 controls)		
Total serum 25(OH)D[a]		
Quartile 1 (6.4-25.3 ng/mL)	1.00	Referent
Quartile 2 (25.4-31.8 ng/mL)	1.04	0.78, 1.39
Quartile 3 (31.9-38.5 ng/mL)	0.94	0.70, 1.26
Quartile 4 (38.6-73.9 ng/mL)	0.71	0.52, 0.98
Continuous 25(OH)D (per 10 ng/mL)	0.87	0.78, 0.98
Regular vitamin D supplement use (≥4 times/week)	0.87	0.70, 1.07
Model 2: Secondary measure of 25(OH)D, Baseline covariates (820 cases, 764 controls)		
Total serum 25(OH)D[b]		
Quartile 1 (9.5-30.8 ng/mL)	1.00	Referent
Quartile 2 (30.9-39.1 ng/mL)	1.41	1.04, 1.91
Quartile 3 (39.2-47.9 ng/mL)	1.57	1.16, 2.13
Quartile 4 (48.0.-137.6 ng/mL)	1.85	1.36, 2.51
Continuous 25(OH)D (per 10 ng/mL)	1.17	1.08, 1.26
Regular vitamin D supplement use (≥4 times/week)	1.39	1.07, 1.81
Model 3: Secondary measure of 25(OH)D, Covariates assessed at time of second blood draw (822 cases, 765 controls)		
Total serum 25(OH)D[c]		
Quartile 1 (9.5-30.8 ng/mL)	1.00	Referent
Quartile 2 (30.9-39.1 ng/mL)	1.50	1.10, 2.04
Quartile 3 (39.2-47.9 ng/mL)	1.55	1.14, 2.11
Quartile 4 (48.0.-137.6 ng/mL)	1.88	1.38, 2.57
Continuous 25(OH)D (per 10 ng/mL)	1.17	1.08, 1.26
Regular vitamin D supplement use (≥4 times/week)	1.41	1.08, 1.85

[a]Adjusted for age at blood draw, race/ethnicity, education, current hormonal birth control use, current hormone therapy use and type, physical activity, BMI, alcohol use, osteoporosis, parity, menopausal status and menopausal status by BMI interaction term

[b]Same covariates as model 1, but used age at second blood draw

[c]Same covariates as model 1 (assessed at time of second blood draw), excluding current hormonal birth control use and grouping two types of hormone therapy together