Circulating vitamin D and breast cancer risk: an international pooling project of 17 cohorts

Visvanathan, K., Mondul, A. M., Zeleniuch-Jacquotte, A., Wang, M., Gail, M. H., Yaun, S.-S., Weinstein, S. J., McCullough, M. L., Eliassen, A. H., Cook, N. R., Agnoli, C., Almquist, M., Buring, J. E., Chen, C., Clendenen, T., Dossus, L., Fedirko, V., Gierach, G. L., Giovannucci, E. L., ... Vitamin D and Breast Cancer Pooling Project (2023). Circulating vitamin D and breast cancer risk: an international pooling project of 17 cohorts. *European Journal of Epidemiology*. Advance online publication. https://doi.org/10.1007/s10654-022-00921-1

Published in:
European Journal of Epidemiology

Document Version:
Peer reviewed version

Queen's University Belfast - Research Portal:
Link to publication record in Queen's University Belfast Research Portal

Publisher rights
Copyright © 2023, This is a U.S. Government work and not under copyright protection in the US; foreign copyright protection may apply.
This work is made available online in accordance with the publisher's policies. Please refer to any applicable terms of use of the publisher.

General rights
Copyright for the publications made accessible via the Queen's University Belfast Research Portal is retained by the author(s) and / or other copyright owners and it is a condition of accessing these publications that users recognise and abide by the legal requirements associated with these rights.

Take down policy
The Research Portal is Queen's institutional repository that provides access to Queen's research output. Every effort has been made to ensure that content in the Research Portal does not infringe any person's rights, or applicable UK laws. If you discover content in the Research Portal that you believe breaches copyright or violates any law, please contact openaccess@qub.ac.uk.

Open Access
This research has been made openly available by Queen's academics and its Open Research team. We would love to hear how access to this research benefits you. – Share your feedback with us: http://go.qub.ac.uk/oa-feedback

Download date:19. Jan. 2025

Circulating vitamin D and breast cancer risk: an international pooling project of 17 cohorts

Kala Visvanathan[1,a], Alison M. Mondul[2,a], Anne Zeleniuch-Jacquotte[3,a], Molin Wang[4,5,6], Mitchell H. Gail[7], Shiaw-Shyuan Yaun[8], Stephanie J. Weinstein[7], Marjorie L. McCullough[9], A. Heather Eliassen[5,6,8], Nancy R. Cook[5,10], Claudia Agnoli[11], Martin Almquist[12], Amanda Black[7], Julie E. Buring[5,10], Chu Chen[13], Yu Chen[3], Tess Clendenen[3], Laure Dossus[14], Veronika Fedirko[15], Gretchen L. Gierach[7], Edward L. Giovannucci[5,6,8], Gary E. Goodman[13], Marc T. Goodman[16], Pascal Guénel[17], Göran Hallmans[18], Susan E. Hankinson[19], Ronald L. Horst[20,b], Tao Hou[8], Wen-Yi Huang[7], Michael E. Jones[21], Corrine E. Joshu[1], Rudolf Kaaks[22], Vittorio Krogh[11], Tilman Kühn[22,23], Marina Kvaskoff[17], I-Min Lee[5,10], Yahya Mahamat-Saleh[17], Johan Malm[24], Jonas Manjer[25], Gertraud Maskarinec[26], Amy E. Millen[27], Toqir K. Mukhtar[28,29], Marian L. Neuhouser[13], Trude E. Robsahm[30], Minouk J. Schoemaker[21], Sabina Sieri[11], Malin Sund[31], Anthony J. Swerdlow[21,32], Cynthia A. Thomson[33], Giske Ursin[30,34], Jean Wactawski-Wende[27], Ying Wang[9], Lynne R. Wilkens[35], Yujie Wu[4], Emilie Zoltick[36], Walter C. Willett[5,8], Stephanie A. Smith-Warner[5,8,c], Regina G. Ziegler[7,c]

[a] Contributed equally as first authors.
[b] Deceased.
[c] Contributed equally as senior authors.

Corresponding author: Regina G. Ziegler
regina.ziegler@nih.gov

Affiliations

1. Departments of Epidemiology and Oncology, Johns Hopkins Bloomberg School of Public Health and Kimmel Cancer Center, Baltimore, MD, USA.
2. Department of Epidemiology, University of Michigan School of Public Health, Ann Arbor, MI, USA.
3. Departments of Population Health and Environmental Medicine, New York University Grossman School of Medicine, New York, NY, USA.
4. Department of Biostatistics, Harvard T.H. Chan School of Public Health, Boston, MA, USA.
5. Department of Epidemiology, Harvard T.H. Chan School of Public Health, Boston, MA, USA.
6. Channing Division of Network Medicine, Department of Medicine, Harvard Medical School, Boston, MA, USA.
7. Division of Cancer Epidemiology and Genetics, National Cancer Institute, National Institutes of Health, Bethesda, MD, USA.
8. Department of Nutrition, Harvard T.H. Chan School of Public Health, Boston, MA, USA.
9. Department of Population Science, American Cancer Society, Atlanta, GA, USA.
10. Division of Preventive Medicine, Brigham and Women's Hospital, Harvard Medical School, Boston, MA, USA.
11. Epidemiology and Prevention Unit, IRCCS National Cancer Institute Foundation, Milan, Italy.
12. Department of Surgery, Skane University Hospital, Lund, Sweden.
13. Division of Public Health Sciences, Fred Hutchinson Cancer Research Center, Seattle, WA, USA.
14. Nutrition and Metabolism Branch, International Agency for Research on Cancer, World Health Organization, Lyon, France.
15. Department of Epidemiology, Rollins School of Public Health and Winship Cancer Institute, Emory University, Atlanta, GA, USA.
16. Cancer Prevention and Control Research Program, Cedars Sinai Samuel Oschin Comprehensive Cancer Institute, Los Angeles, CA, USA.
17. Center for Research in Epidemiology and Population Health (CESP), French National Institute of Health and Medical Research (INSERM), University Paris-Saclay, Villejuif, France.
18. Department of Public Health and Clinical Medicine, Umeå University, Umeå, Sweden.
19. Department of Biostatistics and Epidemiology, School of Public Health and Health Sciences, University of Massachusetts Amherst, Amherst, MA, USA.
20. Heartland Assays, LLC, Ames, IA, USA.
21. Division of Genetics and Epidemiology, The Institute of Cancer Research, London, UK.
22. Division of Cancer Epidemiology, German Cancer Research Center (DKFZ), Heidelberg, Germany.
23. Institute for Global Food Security, Queen's University, Belfast, Northern Ireland.
24. Department of Translational Medicine, Lund University, Malmö, Sweden.
25. Department of Surgery, Skane University Hospital, Lund University, Malmö, Sweden.
26. Cancer Epidemiology Program, University of Hawai'i Cancer Center, Honolulu, HI, USA.
27. Department of Epidemiology and Environmental Health, University at Buffalo, Buffalo, NY, USA.

28. Department of Primary Care and Public Health, Imperial College, London, UK.
29. Department of Public Health and Primary Care, University of Cambridge, Cambridge, UK.
30. Cancer Registry of Norway, Institute of Population-Based Cancer Research, Oslo, Norway.
31. Department of Surgical and Perioperative Sciences, Umeå University, Umeå, Sweden.
32. Division of Breast Cancer Research, The Institute of Cancer Research, London, UK.
33. Department of Health Promotion Sciences, Mel and Enid Zuckerman College of Public Health, University of Arizona and University of Arizona Cancer Center, Tucson, AZ, USA.
34. Department of Nutrition, Institute of Basic Medical Sciences, University of Oslo, Oslo, Norway.
35. Population Sciences in the Pacific, University of Hawai'i Cancer Center, Honolulu, HI, USA.
36. Department of Population Medicine, Harvard Medical School, Boston, MA, USA.

Word count: 4634
2 tables and 5 figures
Supplementary methods: 1922 words
3 supplementary tables

Abstract

Laboratory and animal research support a protective role for vitamin D in breast carcinogenesis, but epidemiologic studies have been inconclusive. To examine comprehensively the relationship of circulating 25-hydroxyvitamin D [25(OH)D] to subsequent breast cancer incidence, we harmonized and pooled participant-level data from 10 U.S. and 7 European prospective cohorts. Included were 10,484 invasive breast cancer cases and 12,953 matched controls. Median age (interdecile range) was 57 (42-68) years at blood collection and 63 (49-75) years at breast cancer diagnosis. Prediagnostic circulating 25(OH)D was either newly measured using a widely accepted immunoassay and laboratory or, if previously measured by the cohort, calibrated to this assay to permit using a common metric. Study-specific relative risks (RRs) for season-standardized 25(OH)D concentrations were estimated by conditional logistic regression and combined by random-effects models. Circulating 25(OH)D increased from a median of 22.6 nmol/L in consortium-wide decile 1 to 93.2 nmol/L in decile 10. Breast cancer risk in each decile was not statistically significantly different from risk in decile 5 in models adjusted for breast cancer risk factors, and no trend was apparent (*P*-trend = 0.64). Compared to women with sufficient 25(OH)D based on Institute of Medicine guidelines (50–<62.5 nmol/L), RRs were not statistically significantly different at either low concentrations (<20 nmol/L, 3% of controls) or high concentrations (100-<125 nmol/L, 3%; ≥125 nmol/L, 0.7%). RR per 25 nmol/L increase in 25(OH)D was 0.99 (95% confidence interval [CI]: 0.95-1.03). Associations remained null across subgroups, including those defined by body mass index, physical activity, latitude, and season of blood collection. Although none of the associations by tumor characteristics reached statistical significance, suggestive inverse associations were seen for distant and triple negative tumors. Circulating 25(OH)D, comparably measured in 17 international cohorts and season-standardized, was not related to subsequent incidence of invasive breast cancer over a broad range in vitamin D status.

Keywords: 25-hydroxyvitamin D, biomarker, breast cancer, calibration, pooled analysis, prospective cohort study, vitamin D

Acknowledgements

We thank Ms. Amelia Zhang-Gross for administrative support and Dr. Micah S. Ziegler for figure design and preparation. BBC3 is a project within the National Cancer Institute Cohort Consortium.

Introduction

Breast cancer is the most common cancer among women, both in the U.S.(1) and worldwide.(2) The 2.3 million cases of breast cancer diagnosed globally in 2020 accounted for one in four of female cancer cases.(2) Breast cancer, already the leading cause of cancer mortality among women worldwide, recently became the leading cause among U.S. women.(2)

Compelling evidence from laboratory and animal research suggests that vitamin D can inhibit cancer initiation and progression by regulating multiple pathways, including differentiation, proliferation, invasion, apoptosis, angiogenesis, inflammation, and immunity.(3, 4) Many of these experimental models used mammary tissue or mammary tumors(3-5) and are thus directly relevant to human breast cancer. Vitamin D receptors and CYP27B1 and CYP24A1, the enzymes that synthesize and metabolize 1α,25-dihydroxyvitamin D, the active form of vitamin D, have been identified in normal human breast tissue and in breast cancers.(5) In addition, vitamin D can modulate the influence of endogenous estrogen, an accepted breast cancer promoter, by decreasing aromatase expression and estrogen production in breast adipose tissue and downregulating estrogen receptor (ER)-alpha expression.(4) Circulating concentrations of 25(OH)D, the accepted measure of vitamin D status, can be increased with minimal risk by dietary supplements, fortified foods, oily fish consumption, and modest sunlight exposure. Therefore, it has been proposed that vitamin D offers a simple, effective strategy for breast cancer prevention.

However, evidence from epidemiologic research, both observational studies and clinical trials, for a role of vitamin D in breast cancer prevention has been inconclusive. Of the 12 cohort studies of prediagnostic circulating 25(OH)D and risk of breast cancer completed at the time we designed our pooling project, two had reported statistically significant inverse associations;(6, 7) five had reported statistically non-significant inverse associations;(8-12) two had reported statistically non-significant positive associations;(13, 14) and three had reported essentially null associations.(11, 15, 16) Importantly, these studies used different assays and laboratories to measure vitamin D, and vitamin D concentrations differ noticeably when different assay methods are employed.(17, 18) Therefore, these studies, and meta-analyses combining them,(19-21) were unable to explore the vitamin D-breast cancer relationship with a common metric.

In the Women's Health Initiative (WHI), which included the largest randomized, placebo-controlled trial of supplemental vitamin D completed to date, 400 IU of vitamin D_3 and 1000 mg of calcium daily did not lower breast cancer incidence.(22, 23) However, the low vitamin D dose, limited adherence, and participants continuing their personal vitamin D supplements during the trial complicated interpretation of results.(22, 24)

In 2011, the Institute for Medicine (IOM; now the National Academy of Medicine) concluded that the evidence of vitamin D benefits for cancer prevention was too inconsistent and insufficient to inform the nutritional requirements for vitamin D, and more research was needed.(3, 25) Thus, IOM public health recommendations for vitamin D are based on bone health. In 2018, the World Cancer Research Fund/American Institute for Cancer Research reported that evidence about vitamin D and breast cancer risk was too limited to reach a conclusion.(26)

We have explored how circulating 25(OH)D is related to subsequent breast cancer incidence in 17 prospective cohorts participating in the international Circulating Biomarkers and Breast and Colorectal Cancer Consortium (BBC3). Participant-level data from each cohort were harmonized and pooled centrally. Circulating 25(OH)D was either newly measured using a single, widely accepted assay and laboratory or, if previously measured by the individual cohort, calibrated to this assay. Advantages of our approach include 1) a wide range in vitamin D status due to diversity across cohort populations, 2) calibration of all 25(OH)D data to a single metric, and 3) a sufficiently large sample size (almost 10,500 breast cancer cases and 13,000 controls) to examine risk in population subgoups and by tumor characteristics.

Methods

Study design

Our study population included 17 prospective cohorts in the U.S. and Europe (**Table 1**).(6-11, 13-16, 27-32) To participate, a cohort needed ≥100 women who developed breast cancer during follow-up and had either prediagnostic 25(OH)D measurements or stored prediagnostic blood samples available for 25(OH)D assays. Each cohort and the BBC3 coordinating center received approval from its institutional review board.

Cases were defined as women with no history of cancer at study entry who then developed primary invasive breast cancer (International Classification of Diseases [ICD] 9 174.0-174.9; ICD 10 C50.0-C50.9) during follow-up. Cases were identified by the individual cohorts using various combinations of methods depending on the cohort: linkage with cancer and/or mortality registries, review of medical and/or health insurance records, self-report, and follow-up with physicians or next-of-kin. Over 98% of cases were confirmed histologically.

Sixteen cohorts utilized a nested case-control design. One or two controls free of breast cancer at the time of diagnosis of the matched case were individually matched to cases on age at blood collection, usually date of blood collection, and other study-specific factors (**Supplementary Table S1**). In PLCO, non-cases were frequency matched to cases on age and date at blood collection.(15)

Measurement of circulating 25(OH)D

Our primary exposure was prediagnostic circulating 25(OH)D measured in plasma or serum. For the case and control(s) in each matched set, the blood samples used for vitamin D measurement were collected before the case was diagnosed with breast cancer. For the five breast cancer studies that had not previously quantified circulating 25(OH)D (**Table 1**), we measured 25(OH)D at Heartland Assays (Ames, IA) using a direct, competitive chemiluminescence immunoassay.(33, 34) For each of the 12 studies that had previously quantified 25(OH)D (**Table 1**), we calibrated its 25(OH)D measurements to the Heartland Assays chemiluminescence immunoassay to correct for assay and laboratory differences in 25(OH)D measurement. For each calibrated study, blood samples from 29 controls, selected uniformly across deciles of the 25(OH)D distribution in study controls, were re-assayed at Heartland Assays (**Supplementary Methods**). A study-specific M-estimation robust linear regression model(35) was used to convert its 25(OH)D data to the concentrations expected had the chemiluminescence

immunoassay been used.(36, 37) Unless otherwise noted, we present calibrated 25(OH)D values for these 12 studies.

Two types of blinded quality controls were randomly inserted into each assay batch by the individual cohorts: 1) Standard Reference Materials with known 25(OH)D concentrations prepared by the National Institute of Standards and Technology (NIST) and 2) cohort-specific quality controls chosen by the cohort (**Supplementary Methods**). Coefficients of variation (CVs) for the chemiluminescence immunoassay, including within- and between-batch variability across both newly assayed and calibrated studies, were 13.4%, 7.8%, and 7.8%, respectively, for NIST reference materials at 17.7, 32.3, and 49.8 nmol/L concentrations. CV's, including within- and between-batch variability, ranged from 5.3% to 14.7% for the cohort-specific quality controls inserted into the newly assayed studies.

To adjust for seasonal variation in circulating 25(OH)D, we calculated a season-standardized 25(OH)D value for each participant, which represents her circulating 25(OH)D averaged over the entire year, according to the weekly 25(OH)D variation observed in her study population. For each study, we used sine-cosine functions and data from controls to model 25(OH)D concentrations as a function of week of blood collection.(37, 38) Then for each participant, the residual, i.e., the difference between her observed 25(OH)D concentration and the concentration predicted by the study-specific model for the week in which her blood was collected, was added to the predicted all-year-round mean 25(OH)D for the controls in her study. All 25(OH)D concentrations presented are season-standardized unless otherwise noted.

Harmonization of covariates and breast tumor characteristics

Each cohort provided participant-level data on demographic, lifestyle, menstrual, reproductive, and medical risk factors, based on questionnaires completed by study participants close to the time of blood collection. At the BBC3 coordinating center, covariates were harmonized using uniform definitions across studies.

Each cohort provided information on the tumor characteristics of its individual cases; these data were harmonized at the BBC3 coordinating center (**Supplementary Methods**).

Statistical analysis

We used a two-stage approach to estimate pooled incidence rate ratios, hereafter called RRs, in most analyses. Study-specific RRs for each nested case-control study were estimated by conditional logistic regression(39) and combined using random-effects models.(40) Heterogeneity in results across studies was evaluated with the Q statistic.(40, 41) For selected analyses, we aggregated the data from all studies into a single dataset (**Supplementary Methods**).

We modeled circulating 25(OH)D in three ways: 1) consortium-wide deciles based on the 25(OH)D distribution in controls from all studies combined, 2) categories suggested by the IOM recommendations for vitamin D(3), and 3) a continuous variable. All three approaches assume that 25(OH)D concentrations from the individual studies use the same metric, and thus require calibration of previously measured 25(OH)D data to the assay chosen for the newly measured studies. Wald tests for trend across consortium-wide deciles or IOM-based categories assigned the median 25(OH)D in

controls to each decile or category. Prior to analyzing 25(OH)D on a continuous scale, we confirmed with restricted cubic spline analyses(42, 43) in the aggregated dataset that the 25(OH)D-breast cancer association was consistent with linearity (*P*-nonlinearity > 0.05).

For two-stage analyses by consortium-wide deciles, we selected decile 5 as the referent since it permitted us to estimate RRs at low and high concentrations and included cases and controls from each of the participating studies. In order to compare risk between extreme deciles, we combined all studies into a single, aggregated dataset since one study did not contribute cases or controls to decile 1.

Three models were used to control for confounding. Model 1 conditioned on study-specific matching factors, which included age at blood collection and usually date of blood collection (**Supplementary Table S1**). Alternatively, in unconditional models, age at blood collection, elapsed time from blood collection to assay, and for EPIC and MEC, study center were included as covariates in Model 1. Model 2 added to model 1 body mass index (BMI) and physical activity. For BMI, a variable combining BMI categories with menopausal status at date of blood collection was used because breast cancer risk is inversely associated with premenopausal BMI and positively associated with postmenopausal BMI.(44) Model 3, the model presented unless otherwise noted, was further adjusted for established breast cancer risk factors: race, family history of breast cancer, age at menarche, parity/age at first birth, oral contraceptive use, menopausal hormone therapy, and alcohol intake (**Supplementary Methods**).

We explored whether the association of circulating 25(OH)D with breast cancer risk varied across population subgroups or by tumor characteristics (**Supplementary Methods).**

All *P* values were based on two-sided tests and considered statistically significant if $P < 0.05$. Analyses were conducted using SAS version 9.4 (Cary, NC).

Results

Our study population included 10,484 female breast cancer cases and 12,953 controls from 10 US and 7 European cohorts (**Table 1**). Median (interdecile range) age was 57 (42-68) years at blood collection, and 63 (49-75) years at breast cancer diagnosis. Median (interdecile range) elapsed time from blood collection to diagnosis was 5.1 (1-13) years. 91% of participants were White.

For five studies, contributing 14% of participants (1595 cases, 1595 controls), circulating 25(OH)D was newly measured at Heartland Assays (**Table 1**). For the 12 calibrated studies, which had previously measured 25(OH)D, the change after calibration in median 25(OH)D concentrations among controls ranged from -31 nmol/L to +10 nmol/L (**Figure 1**). Vitamin D status, based on season-standardized, calibrated data, differed markedly among the participating cohorts, with median 25(OH)D concentrations increasing from 31 to 71 nmol/L across studies (**Table 1; Figure 1**).

Breast cancer risk by consortium-wide 25(OH)D deciles

Across consortium-wide deciles, 25(OH)D concentrations increased 4-fold, from a median of 22.6 nmol/L in decile 1 to 93.2 nmol/L in decile 10 (**Table 2**). Breast cancer risk in each decile was not statistically significantly different from that in decile 5 for all three models (RR range = 0.96-1.11; all *P*-trend ≥ 0.20). Comparing women in the

highest decile of circulating 25(OH)D to those in the lowest, the RR, fully adjusted for breast cancer risk factors, was 0.98 (95% CI: 0.85-1.12).

Confounding by breast cancer risk factors was minimal; results for the fully adjusted model (model 3) were very similar to those for the two simpler models (**Table 2**). No statistically significant differences in results were observed across the 17 studies (*P*-heterogeneity = 0.67 and 0.52 for deciles 1 and 10, respectively). Excluding one study at a time revealed that no single study substantially influenced the results (data not shown).

Breast cancer risk by 25(OH)D categories based on IOM clinical guidance

We examined how breast cancer risk was related to the vitamin D guidelines established by IOM.(3) Relative to women with sufficient 25(OH)D (50-<75 nmol/L; 36% of consortium controls), RRs were 0.93 (95% CI: 0.83-1.04) for women with deficient levels (<30 nmol/L; 13%) and 1.03 (95% CI: 0.96-1.11) for insufficient levels (30-<50 nmol/L; 34%). For 25(OH)D concentrations beyond sufficient (>75 nmol/L; 16%), RR was 0.95 (95% CI: 0.87-1.03).

We also evaluated finer categories to assess breast cancer risk at very low and very high 25(OH)D concentrations (**Figure 2; Supplementary Table S2**). Compared with women with 25(OH)D in the lower range of sufficiency (50-<62.5 nmol/L; 21% of consortium controls), statistically non-significant RRs were noted for all 25(OH)D categories, including <20 nmol/L (3%), 100-<125 nmol/L (3%), and ≥125 nmol/L (0.7%). No statistically significant differences in results were observed across the 17 studies (*P*-heterogeneity = 0.57 and 0.25 for <20 and ≥125 nmol/L, respectively).

Breast cancer risk by continuous 25(OH)D

The RR for a 25 nmol/L increase in circulating 25(OH)D was 0.99 (95% CI: 0.95-1.03) (**Table 2**). Cohort-specific RRs ranged from 0.89 to 1.16, and none reached statistical significance (*P*-heterogeneity across the 17 studies = 0.63) (**Figure 3**).

Circulating 25(OH)D-breast cancer associations in population subgroups and by tumor characteristics

The essentially null association between circulating 25(OH)D, modeled on a continuous scale, and breast cancer risk did not vary across subgroups defined by BMI, physical activity, family history of breast cancer, menopausal hormone therapy, latitude, or geographic region; all *P*-interaction > 0.05 (**Figure 4)**. In addition, results were not statistically significantly different for participants donating blood in the winter vs. summer (*P*-interaction = 0.83) or for studies in which 25(OH)D was newly measured vs. calibrated (*P*-interaction = 0.94). Differences by race were not statistically significant (*P*-heterogeneity = 0.90). For a 25 nmol/L increase in 25(OH)D, RR was 0.98 (95% CI: 0.95-1.02) in Whites (9,579 cases); 1.28 (95% CI: 0.99-1.65) in Blacks (290 cases); and 1.13 (95% CI: 0.76-1.68) in Asians (275 cases).

Substantial numbers of breast cancer subtypes were available for analysis. For example, among the 9,787 breast cancer cases with morphology data (93% of all cases), 7,148 (73%) were ductal only and 1,179 (12%) were lobular only. Of the 8,788 cases with ER data (84% of all cases), 6,949 (79%) were ER-positive (ER+); of the 8,295 cases with PR data (79% of all cases), 5,444 (66%) were PR-positive (PR+). No

statistically significant associations with circulating 25(OH)D, modelled on a continuous scale, were observed for any of the tumor subtypes we examined, including those defined by stage, grade, morphology, receptor status, or molecular features(45) (**Figure 4; Supplementary Methods**). However, a 17% reduction in risk (RR per 25 nmol/L increase = 0.83; 95% CI: 0.68-1.02) was noted for distant disease (stage 4; 319 cases); and an 11% reduction (RR = 0.89; 95% CI: 0.72-1.08), for triple negative disease (ER−/progesterone receptor [PR]−/human epidermal growth factor 2 [HER2]−; 365 cases). Exclusion of cases diagnosed ≤2 years after blood collection and their matched controls did not attenuate either association (distant disease: RR = 0.77; 95% CI: 0.60-0.98; 269 cases; triple negative disease: RR = 0.87; 95% CI: 0.70-1.08; 310 cases). These two suggestive inverse associations were independent, as only 2% of the women with triple negative disease also presented with distant disease. Of note, risk was not reduced for ER−/PR− or ER−/PR−/HER2+ (HER2-enriched) disease.

We did not observe differences by age at diagnosis (*P*-heterogeneity = 0.13) (**Figure 4**). We then evaluated risk of early onset breast cancer (diagnosed <50 y; 1,212 cases; 12%) and late onset breast cancer (diagnosed ≥75 y; 1,045 cases; 10%) across consortium-wide 25(OH)D deciles. For each subtype, none of the RRs across deciles, relative to decile 5, were statistically significant (data not shown); *P*-trend was 0.77 for early onset and 0.19 for late onset disease.

Breast cancer risk was not related to elapsed time from blood collection to diagnosis (**Figure 4)**, which suggests that preclinical disease did not influence our findings**.**

Influence of the season of blood collection

To extend our analysis of the influence of season of blood collection, we explored the 25(OH)D-breast cancer relationship among women who donated blood in winter (defined as weeks 2-14) or summer (defined as weeks 28-40) (**Supplementary Methods**) using season-specific consortium-wide deciles of circulating 25(OH)D (**Figure 5**). Median (interdecile range) 25(OH)D was 45 (22-75) and 61 (35-94) nmol/L among controls donating in winter and summer, respectively (**Figure 5**). There was no apparent change in breast cancer risk over the 25(OH)D ranges observed in each season (*P*-trend = 0.72 and 0.95 for winter and summer ranges, respectively).

Sensitivity analyses

Results were similar using a two-stage approach with either random-effects or fixed-effects models in the second stage and using an aggregated dataset (**Supplementary Table S3**). Finer control of BMI for all studies and of physical activity for the 11 studies with sufficiently detailed information did not noticeably change RRs (data not shown). To minimize including women with preclinical disease, we excluded cases diagnosed within the first two or five years after blood collection (2,049 and 5,124 cases, respectively) and their matched controls; our results were essentially unchanged (data not shown). All sensitivity analyses were performed for 25(OH)D modeled as consortium-wide deciles, IOM-based categories, and a continuous measure.

Discussion

Across a wide range in vitamin D status, we saw no evidence that circulating

25(OH)D concentrations were associated with subsequent incidence of invasive breast cancer overall. Breast cancer risk did not differ across consortium-wide deciles of 25(OH)D, even though median 25(OH)D concentrations quadrupled from decile 1 (22.6 nmol/L) to decile 10 (93.2 nmol/L). Relative to women with sufficient vitamin D (50-<75 nmol/L), based on current IOM guidance,(3, 25), breast cancer risk was not different in women considered deficient (<30 nmol/L) or insufficient (30-<50 nmol/L). Even at high circulating 25(OH)D (100-<125 nmol/L and ≥125 nmol/L), risk was unchanged. Breast cancer risk for a 25 nmol/L increase in 25(OH)D was 0.99 (95% CI: 0.95-1.03).

Although the participating cohorts differed by geography, demographics, lifestyle, and diet, results were generally similar across the 17 cohorts. There was no evidence of statistically significant heterogeneity in the results across the studies whether 25(OH)D was modeled as consortium-wide deciles, IOM-based cutpoints, or a continuous variable. In addition, eliminating any single study did not substantially alter results.

Several prospective studies of circulating 25(OH)D and breast cancer did not find statistically significant associations overall but noted relationships limited to subgroups. Statistically significant inverse associations have been reported in younger women,(7, 11) for premenopausal breast cancer,(46) in menopausal hormone therapy users,(14) in leaner women,(47) and at higher latitudes(16) while statistically significant positive associations have been reported in heavier women.(13, 14, 47) By calibrating 25(OH)D data from multiple studies to a single, widely accepted assay and harmonizing and pooling covariate and outcome data, we could generate more robust RR estimates for population subgroups than individual studies. With almost 10,500 breast cancer cases, more than seven times the number included in the largest study of prediagnostic 25(OH)D and breast cancer published to date,(48) we found null associations across all the subgroups we examined, including those defined by BMI, latitude, and season of blood collection. However, with <300 cases among Blacks or Asians, we had limited power to evaluate risk by race.

Similarly, most studies of circulating 25(OH)D and breast cancer have had too few cases to generate robust RR estimates by tumor characteristics. We found no statistically significant associations by stage, grade, morphology, receptor status, or age at onset. However, we did observe suggestive, independent inverse associations for two breast cancer subtypes with poor prognoses: distant disease (319 cases) and triple negative disease (365 cases). It is unlikely that preclinical disease underlies these findings since each persisted when cases diagnosed ≤2 years after blood collection and their matched controls were excluded from analysis.

Circulating 25(OH)D fluctuates seasonally since sunlight stimulates its endogenous production. Among consortium controls donating blood in the winter, 24% were deficient (<30 nmol/L) for vitamin D, but only 6% of controls donating in the summer were deficient. It has been hypothesized that women whose circulating 25(OH)D drops to especially low levels in the winter or whose 25(OH)D fails to rise sufficiently in the summer to replenish adipose tissue stores might be at increased breast cancer risk. Two cohorts in our consortium have reported season-specific effects. In NHS, 25(OH)D measured in summer samples (collected May–October) was statistically significantly inversely associated with breast cancer risk, but winter levels (November-April) were not.(48) Similar, but statistically non-significant, findings were

reported by NYU WHS.(11) However, in our season-specific analyses, we saw no evidence that 25(OH)D measured during the summer or winter was related to breast cancer risk.

At its initiation, our consortium included nearly all completed prospective studies of circulating 25(OH)D and invasive breast cancer risk (12 cohorts, 8,889 cases), and then increased case numbers by 18% by assaying 25(OH)D in five additional cohorts (1,595 cases). Of the recently published prospective studies of invasive breast cancer not included in our pooling project, six, conducted in Germany (137 cases),(49) Denmark (159 cases),(50) France (233 cases),(47) Japan (239 cases),(46) southeastern Australia (634 cases),(51) and Turkey (57 cases),(52) have reported null associations, consistent with our findings. Three, two in the US (1611(53) and 77(54) cases) and one in western Australia (113 cases),(55) have reported statistically significant inverse associations; and one, in Europe (378 cases),(56) a statistically significant positive association. The largest study to date of prediagnostic circulating 25(OH)D and risk of *in situ* ductal breast cancer (UK; 1340 cases)(57) found no relationship, similar to our study of invasive disease. None of these recent studies have the range in vitamin D status achieved by our international pooling project. In addition, the number of invasive breast cancer cases in all the recent studies combined is less than 35% of the cases included in our pooling project.

Although a number of randomized, placebo-controlled trials of vitamin D supplementation have been published,(58-60) few were sufficiently large to evaluate breast cancer incidence. The Vitamin D and Omega-3 Trial (VITAL) is the largest and longest randomized, placebo-controlled trial of moderate-to-high dose vitamin D supplementation to date.(61) More than 13,000 women, ≥55 years, were given 2000 IU/day of supplemental vitamin D_3 (cholecalciferol) or placebo and followed for a median of 5.3 years. No statistically significant differences between groups were observed for invasive breast cancer (hazard ratio [HR] = 1.02; 95% CI: 0.79-1.31). Our null results for invasive breast cancer overall concur with and extend these findings. VITAL had limited power to examine breast cancer, with only 246 invasive cases(61) and 10 advanced (metastatic or fatal) cases(62). Our pooling project, with much larger numbers of breast cancers, had more statistical power to explore breast cancer risk overall, by tumor characteristics, and in population subgroups. For example, VITAL noted that vitamin D statistically significantly reduced the incidence of all cancers combined in normal-weight individuals (BMI < 25 kg/m^2).(59, 61) However, in our analyses of breast cancer, we found no interaction with BMI.

In our pooling project, we were able to evaluate lower 25(OH)D concentrations than VITAL. Only 13% of VITAL participants had a baseline 25(OH)D <50 nmol/L,(61) the level required for vitamin D sufficiency by IOM guidelines, whereas 48% of our consortium controls were below sufficiency. Conversely, the 25(OH)D concentrations attained through supplementation in VITAL, averaging ≥102.5 nmol/L,(59) were somewhat higher than in our consortium. Less than 4% of our controls had 25(OH)D concentrations ≥100 nmol/L.

Recently VITAL reported that advanced cancers (metastatic or fatal) were statistically significantly reduced among those randomized to vitamin D compared to placebo (HR = 0.83; 95% CI: 0.69-0.99),(62) even though all invasive cancers combined were not statistically significantly reduced. Our suggestive protective effects for distant

and triple negative breast cancer, subtypes with poor prognoses, concur with these VITAL results.

Our generally null results also agree with those of the WHI vitamin D-calcium trial, the largest trial of low-dose vitamin D supplementation.(22) Over 36,000 postmenopausal women were given 400 IU of vitamin D_3 plus 1000 mg of calcium daily or placebo for a mean of 7.0 years. No statistically significant differences between groups were observed for invasive breast cancer at the end of the trial (HR = 0.96; 95% CI: 0.85-1.09; 1,074 cases)(22) or after five additional years of followup (HR = 1.04; 95% CI: 0.94-1.14; 1,667 cases).(23)

A critical strength of our pooling project is how we addressed the challenges that arise when combining circulating biomarker data from multiple studies. There are often substantial differences in accuracy among assays and laboratories in measuring 25(OH)D(17, 18), as well as other circulating biomarkers. Calibrating previously measured 25(OH)D readings to the widely accepted assay and laboratory chosen for our new 25(OH)D measurements enabled us to use absolute 25(OH)D concentrations, rather than study-specific quantiles, as our metric and confidently estimate RR across the wide range of vitamin D exposure provided by our 17 international cohorts. In addition, to adjust for blood samples being collected throughout the year, we standardized the 25(OH)D value for each woman to her estimated all-year-round average 25(OH)D, based on models of the weekly 25(OH)D changes in her own cohort.

Additional strengths of our study include its size and scope --- almost 10,500 breast cancer cases and 13,000 matched controls from 17 international prospective cohorts. Followup was longer than in most randomized trials. Across our studies, median time from blood collection to diagnosis ranged from 1.7 to 12.4 years, and exceeded 5 years in nine studies. Only studies that collected blood samples before breast cancer diagnosis were included, which reduces the possibility that disease, its diagnosis, or its treatment altered circulating 25(OH)D. We harmonized participant-level data on exposures and outcomes and used a common statistical methodology for analyses. This approach reduces potential heterogeneity across individual studies and facilitates better control for confounding than meta-analyses that combine results, not participant-level data, from published studies. Of note, our consortium's study of circulating 25(OH)D and colorectal cancer, which with 5,706 cases was approximately half the size of our breast cancer study, found a statistically significant inverse association overall, as well as inverse associations in all subgroups.(63) Since the two studies shared the same methodology, same assay, and many of the same cohorts, this comparison suggests our breast cancer study would have seen an association with 25(OH)D if one existed.

We also considered potential limitations of our study. Excess weight and physical inactivity are breast cancer risk factors and inversely correlated with circulating 25(OH)D.(64) Adjustment for both variables did not noticeably change our results, nor did adjustment for other breast cancer risk factors. Although we cannot completely eliminate the possibility of residual confounding, the minimal confounding observed increases our confidence in the validity of our findings. In addition, our results are based on a single blood collection, at a median of 5.1 years before diagnosis, which is an imperfect measure of long-term vitamin D status. However, studies of repeat measures of 25(OH)D have reported relatively stable readings over time, with intraclass

correlation coefficients of 0.7 for blood samples collected 3-5 years apart and of 0.5–0.6 for samples collected approximately 10 years apart.(11, 48, 65, 66) Finally, we had limited power to examine associations in non-White populations.

In summary, in our large international cohort consortium, circulating 25(OH)D, measured comparably in 17 cohorts by calibrating to a single assay, was not related to subsequent incidence of invasive breast cancer across a wide range of concentrations (<20 to ≥100 nmol/L). No statistically significant associations were noted within population subgroups, by tumor characteristics or molecular subtypes, or for winter or summer measurements. Suggestive inverse associations with both triple negative and distant breast cancer, subtypes with poor prognoses, merit further investigation. Our results do not support adoption of routine vitamin D measurement or treatment of low 25(OH)D levels for breast cancer prevention.

Statements and Declarations

Author contributions

The Circulating Biomarkers and Breast and Colorectal Cancer Consortium was conceived by WCW, SSW, and RGZ and directed by SSW and RGZ. KV, AMM, AZJ, MW, MHG, SSY, SJW, MLM, AHE, NRC, TH, SSW, and RGZ designed the breast cancer study and developed the methodology. Data were provided by investigators from each participating cohort and harmonized by KV, AMM, AZJ, SSY, TH, TKM, SSW, and RGZ. The vitamin D assays were conducted in the laboratory of RLH. Epidemiologic and statistical analyses were conducted by MW, MHG, SSY, and TH with KV, AMM, AZJ, SSW, and RGZ providing input and interpretation. RGZ wrote the initial draft of the manuscript, with KV, AMM, AZJ, and SSW providing extensive input and review on multiple versions of the manuscript. All authors reviewed, edited, and approved the final version of the manuscript.

Funding for consortium

The centralization, checking, harmonization, and statistical analyses of the participant level data from the cohorts; the 25(OH)D assays required for newly measured studies and calibration of previously measured studies; and the laboratory quality control procedures were supported by the National Cancer Institute, National Institutes of Health (R01 CA152071); the Intramural Research Program of the Division of Cancer Epidemiology and Genetics, National Cancer Institute, National Institutes of Health; and the Breast Cancer Research Foundation.

Funding, acknowledgements, and disclaimers for individual cohorts

<u>Beta-Carotene and Retinol Efficacy Trial (CARET)</u>

This work was supported by the National Institutes of Health (U01 CA063673, UM1 CA167462, and UO1 CA167462).

<u>Breakthrough Generations Study (BGS)</u>

We thank Breast Cancer Now and the Institute of Cancer Research for support and funding of the Breakthrough Generations Study, now called the Generations Study. We acknowledge National Health Service funding to the Royal Marsden/ICR NIHR Biomedical Research Centre. We thank the study participants; study staff; and the doctors, nurses, health care providers, and health information sources who have contributed to the study.

<u>Cancer Prevention Study-II (CPS-II)</u>

The American Cancer Society funds the creation, maintenance, and updating of the CPS-II cohort. We acknowledge the contribution to this study by central cancer registries supported through the Centers for Disease Control and Prevention and the Surveillance, Epidemiology, and End Results program. The views expressed here are those of the authors and do not necessarily represent the American Cancer Society or the American Cancer Society Cancer Action Network.

<u>CLUE II: Campaign Against Cancer and Heart Disease (CLUE II)</u>

CLUE II was supported by the National Cancer Institute of the National Institutes of Health (U01 CA86308, P30 CA006973) and the National Institute on Aging of the National Institutes of Health (U01 AG18033). The collection and availability of cancer registry data was also supported by the Centers for Disease Control and Prevention

(cooperative agreement NU58DP006333). Cancer data was provided by the Maryland Cancer Registry, Center for Cancer Prevention and Control, Maryland Department of Health, with funding from the State of Maryland and the Maryland Cigarette Restitution Fund. We thank the participants of CLUE II and appreciate the continued efforts of the staff at the Johns Hopkins George W. Comstock Center for Public Health Research and Prevention in the conduct of the CLUE Cohort Studies. The contents of the article are solely the responsibility of the authors and do not necessarily represent the official views of the Centers for Disease Control and Prevention or the Department of Health and Human Services.

Etude Epidémiologique auprès de Femmes de l'Education Nationale (E3N)

The E3N cohort from the French National Institute of Health and Medical Research (INSERM) was supported by the Mutuelle Générale de l'Education Nationale, the Gustave Roussy Institute, and the French League against Cancer (LNCC). E3N-E4N is also supported by the French National Research Agency (ANR) under the Investment for the Future Program (PIA) (ANR-10-COHO-0006) and by the French Ministry of Higher Education, Research and Innovation (subsidy for public service charges number 2102 918823).

European Investigation into Cancer and Nutrition (EPIC)

The coordination of EPIC is financially supported by International Agency for Research on Cancer (IARC) and also by the Department of Epidemiology and Biostatistics, School of Public Health, Imperial College London, which has additional infrastructure support provided by the National Institute for Health Research Imperial Biomedical Research Centre (BRC). The national cohorts are supported by the Danish Cancer Society (Denmark); Ligue Contre le Cancer, Institut Gustave Roussy, Mutuelle Générale de l'Education Nationale, and Institut National de la Santé et de la Recherche Médicale (INSERM) (France); German Cancer Aid, German Cancer Research Center (DKFZ), German Institute of Human Nutrition Potsdam-Rehbruecke (DIfE), and Federal Ministry of Education and Research (BMBF) (Germany); Associazione Italiana per la Ricerca sul Cancro (AIRC-Italy), Compagnia di San Paolo, and National Research Council (Italy); Dutch Ministry of Public Health, Welfare and Sports (VWS), Netherlands Cancer Registry (NKR), LK Research Funds, Dutch Prevention Funds, Zorg Onderzoek Nederland (ZON), World Cancer Research Fund (WCRF), and Statistics Netherlands (The Netherlands); Health Research Fund (FIS) - Instituto de Salud Carlos III (ISCIII), Regional Governments of Andalucía, Asturias, Basque Country, Murcia and Navarra, and the Catalan Institute of Oncology - ICO (Spain); Swedish Cancer Society, Swedish Research Council, and County Councils of Skåne and Västerbotten (Sweden); and Cancer Research UK (14136 and C8221/A29017) and Medical Research Council (grant numbers 1000143 and MR/M012190/1) (United Kingdom). Where authors are identified as personnel of the International Agency for Research on Cancer / World Health Organization, the authors alone are responsible for the views expressed in this article and they do not necessarily represent the decisions, policy or views of the International Agency for Research on Cancer / World Health Organization.

Hormones and Diet in the Etiology of Breast Cancer Study (ORDET)

This study was supported by the Italian Association for Cancer Research (AIRC).

Janus Serum Bank (JANUS)

We sincerely thank the Cancer Registry of Norway and the Norwegian Cancer

Society for financial funding.
Malmö Diet and Cancer Study (MDCS)

This work was supported by the Swedish Cancer Society.

Multiethnic Cohort Study (MEC)

This work was supported by the National Institutes of Health (P01 CA33619 and U01 CA164973).

New York University Women's Health Study (NYU WHS)

This work was supported by the National Institutes of Health (R01 CA098661, U01 CA182934, P30 CA016087, and P30 ES000260).

Northern Sweden Health and Disease Study (NSHDS)

This work was supported by the Swedish Cancer Society, Swedish Research Council, and County Councils of Västerbotten (Sweden), and the National Institutes of Health (R01 CA098661).

Nurses' Health Study (NHS)

This work was supported by the National Institutes of Health (UM1 CA186107, R01 CA49449, and P01 CA87969). The study protocol was approved by the institutional review boards of the Brigham and Women's Hospital and Harvard T.H. Chan School of Public Health, and those of participating registries as required. We would like to thank the following state cancer registries for their help: AL, AZ, AR, CA, CO, CT, DE, FL, GA, ID, IL, IN, IA, KY, LA, ME, MD, MA, MI, NE, NH, NJ, NY, NC, ND, OH, OK, OR, PA, RI, SC, TN, TX, VA, WA, WY. The content is solely the responsibility of the authors and does not necessarily represent the official views of the National Institutes of Health.

Nurses' Health Study II (NHSII)

This work was supported by the National Institutes of Health (U01 CA176726, R01 CA67262, R01 CA50385, and P50 CA089393). The study protocol was approved by the institutional review boards of the Brigham and Women's Hospital and Harvard T.H. Chan School of Public Health, and those of participating registries as required. We would like to thank the following state cancer registries for their help: AL, AZ, AR, CA, CO, CT, DE, FL, GA, ID, IL, IN, IA, KY, LA, ME, MD, MA, MI, NE, NH, NJ, NY, NC, ND, OH, OK, OR, PA, RI, SC, TN, TX, VA, WA, WY. The content is solely the responsibility of the authors and does not necessarily represent the official views of the National Institutes of Health.

Prostate, Lung, Colorectal and Ovarian Cancer Screening Trial (PLCO)

This research was supported by the Intramural Research Program of the Division of Cancer Epidemiology and Genetics and by contracts from the Division of Cancer Prevention, National Cancer Institute, National Institutes of Health. We thank the PLCO screening center investigators and the staff from Information Management Services, Inc. and Westat, Inc. Most importantly, we thank the study participants for their contributions that made this study possible.

Women's Health Initiative (WHI)

The WHI program was supported by the National Heart, Lung, and Blood Institute of the National Institutes of Health, U.S. Department of Health and Human Services (HHSN268201600018C, HHSN268201600001C, HHSN268201600002C, HHSN268201600003C, and HHSN268201600004C).

Women's Health Study (WHS)

This work was supported by the National Institutes of Health (CA047988, CA182913, HL043851, HL080467, and HL099355).

Data availability

The data described in the manuscript will not be made available because we do not have permission to disclose or release the data from the participating cohorts, as specified in their executed data use agreements.

Competing interests

No authors have reported conflicts of interest related to the information presented in this manuscript.

Ethics approval and consent to participate

Each participating cohort and the BBC3 coordinating center received approval of this study from its institutional ethics review board.

Consent to publish

Not applicable.

References

1. Siegel RL, Miller KD, Fuchs HE, Jemal A. Cancer Statistics, 2021. CA Cancer J Clin. 2021;71(1):7-33. doi:10.3322/caac.21654
2. World Health Organization, International Agency for Research on Cancer. Global Cancer Observatory. 2021 [cited 2021]. Available from https://gco.iarc.fr/.
3. Dietary reference intakes for calcium and vitamin D. Washington, D.C.: The National Academies Press; 2011.
4. Feldman D, Krishnan AV, Swami S, Giovannucci E, Feldman BJ. The role of vitamin D in reducing cancer risk and progression. Nat Rev Cancer. 2014;14(5):342-57. doi:10.1038/nrc3691
5. Welsh J. Function of the vitamin D endocrine system in mammary gland and breast cancer. Mol Cell Endocrinol. 2017;453:88-95. doi:10.1016/j.mce.2017.04.026
6. 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. doi:10.1186/1471-2407-14-29
7. 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):2341-50. doi:10.1158/1055-9965.EPI-10-0264
8. 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):1991-7.
9. Almquist M, Bondeson AG, Bondeson L, Malm J, Manjer J. Serum levels of vitamin D, PTH and calcium and breast cancer risk-a prospective nested case-control study. Int J Cancer. 2010;127(9):2159-68. doi:10.1002/ijc.25215
10. Neuhouser ML, Manson JE, Millen A, et al. The influence of health and lifestyle characteristics on the relation of serum 25-hydroxyvitamin D with risk of colorectal and breast cancer in postmenopausal women. Am J Epidemiol. 2012;175(7):673-84. doi:10.1093/aje/kwr350
11. 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(1):R15. doi:10.1186/bcr3390
12. 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 research and treatment. 2012;133(3):1077-88. doi:10.1007/s10549-012-2012-x
13. Eliassen AH, Spiegelman D, Hollis BW, Horst RL, Willett WC, Hankinson SE. Plasma 25-hydroxyvitamin D and risk of breast cancer in the Nurses' Health Study II. Breast Cancer Res. 2011;13(3):R50. doi:10.1186/bcr2880
14. 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):1689-700. doi:10.1002/ijc.28172
15. Freedman DM, Chang SC, 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. 2008;17(4):889-94.

16. 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. 2009;11(4).
17. Sempos CT, Vesper HW, Phinney KW, Thienpont LM, Coates PM, Vitamin DSP. Vitamin D status as an international issue: national surveys and the problem of standardization. Scand J Clin Lab Invest Suppl. 2012;243:32-40. doi:10.3109/00365513.2012.681935
18. Carter GD, Berry J, Durazo-Arvizu R, et al. Hydroxyvitamin D assays: An historical perspective from DEQAS. J Steroid Biochem Mol Biol. 2018;177:30-5. doi:10.1016/j.jsbmb.2017.07.018
19. Estebanez N, Gomez-Acebo I, Palazuelos C, Llorca J, Dierssen-Sotos T. Vitamin D exposure and risk of breast cancer: a meta-analysis. Sci Rep. 2018;8(1):9039. doi:10.1038/s41598-018-27297-1
20. Hossain S, Beydoun MA, Beydoun HA, Chen X, Zonderman AB, Wood RJ. Vitamin D and breast cancer: A systematic review and meta-analysis of observational studies. Clin Nutr ESPEN. 2019;30:170-84. doi:10.1016/j.clnesp.2018.12.085
21. Song D, Deng Y, Liu K, et al. Vitamin D intake, blood vitamin D levels, and the risk of breast cancer: a dose-response meta-analysis of observational studies. Aging (Albany NY). 2019;11(24):12708-32. doi:10.18632/aging.102597
22. 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):1581-91.
23. Cauley JA, Chlebowski RT, Wactawski-Wende J, et al. Calcium plus vitamin D supplementation and health outcomes five years after active intervention ended: the Women's Health Initiative. J Womens Health (Larchmt). 2013;22(11):915-29. doi:10.1089/jwh.2013.4270
24. Bolland MJ, Grey A, Gamble GD, Reid IR. 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):1144-9. doi:10.3945/ajcn.111.015032
25. Ross AC, Manson JE, Abrams SA, et al. The 2011 report on dietary reference intakes for calcium and vitamin D from the Institute of Medicine: what clinicians need to know. J Clin Endocrinol Metab. 2011;96(1):53-8. doi:10.1210/jc.2010-2704
26. World Cancer Research Fund, American Institute for Cancer Research. Continuous Update Project Expert Report 2018 [Internet]. Diet, nutrition, physical activity, and breast cancer. 2018 [cited 2021]. Available from dietandcancerreport.org.
27. Gross AL, Newschaffer CJ, Hoffman-Bolton J, Rifai N, Visvanathan K. Adipocytokines, inflammation, and breast cancer risk in postmenopausal women: a prospective study. Cancer Epidemiol Biomarkers Prev. 2013;22(7):1319-24. doi:10.1158/1055-9965.EPI-12-1444
28. Omenn GS, Goodman G, Thornquist M, et al. The Beta-Carotene and Retinol Efficacy Trial (CARET) for chemoprevention of lung cancer in high risk populations: smokers and asbestos-exposed workers. Cancer Res. 1994;54(7 Suppl):2038s-43s.
29. Sieri S, Muti P, Claudia A, et al. Prospective study on the role of glucose metabolism in breast cancer occurrence. Int J Cancer. 2012;130(4):921-9. doi:10.1002/ijc.26071

30. Swerdlow AJ, Jones ME, Schoemaker MJ, et al. The Breakthrough Generations Study: design of a long-term UK cohort study to investigate breast cancer aetiology. Br J Cancer. 2011;105(7):911-7. doi:10.1038/bjc.2011.337
31. Ridker PM, Cook NR, Lee IM, et al. A randomized trial of low-dose aspirin in the primary prevention of cardiovascular disease in women. N Engl J Med. 2005;352(13):1293-304. doi:10.1056/NEJMoa050613
32. Langseth H, Gislefoss RE, Martinsen JI, Dillner J, Ursin G. Cohort Profile: The Janus Serum Bank Cohort in Norway. Int J Epidemiol. 2017;46(2):403-4g. doi:10.1093/ije/dyw027
33. Hollis BW. Measuring 25-hydroxyvitamin D in a clinical environment: challenges and needs. Am J Clin Nutr. 2008;88(2):507S-10S. doi:10.1093/ajcn/88.2.507S
34. Wagner D, Hanwell HE, Vieth R. An evaluation of automated methods for measurement of serum 25-hydroxyvitamin D. Clin Biochem. 2009;42(15):1549-56. doi:10.1016/j.clinbiochem.2009.07.013
35. Rousseeuw PJ, Leroy AM. Robust regression and outlier detection. New York: Wiley; 1987.
36. Sloan A, Song Y, Gail MH, et al. Design and analysis considerations for combining data from multiple biomarker studies. Statistics in Medicine. 2019;38(8):1303-20. doi:10.1002/sim.8052
37. Gail MH, Wu J, Wang M, et al. Calibration and seasonal adjustment for matched case-control studies of vitamin D and cancer. Statistics in Medicine. 2016;35(13):2133-48. doi:10.1002/sim.6856
38. Bliss CI. Periodic regression in biology and climatology. New Haven,: Connecticut Agricultural Experiment Station; 1958.
39. Prentice RL, Breslow NE. Retrospective studies and failure time models. Biometrika. 1978;65:153-8.
40. DerSimonian R, Laird N. Meta-analysis in clinical trials. Control Clin Trials. 1986;7(3):177-88. doi:10.1016/0197-2456(86)90046-2
41. Cochran WG. The combination of estimates from different experiments. Biometrics. 1954;10(1):101-29.
42. Durrleman S, Simon R. Flexible regression models with cubic splines. Stat Med. 1989;8(5):551-61. doi:10.1002/sim.4780080504
43. Smith PL. Splines as a useful and convenient statistical tool. Am Stat. 1979;33:57-62.
44. van den Brandt PA, Spiegelman D, Yaun SS, et al. Pooled analysis of prospective cohort studies on height, weight, and breast cancer risk. American Journal of Epidemiology. 2000;152(6):514-27. doi:10.1093/aje/152.6.514
45. Goldhirsch A, Winer EP, Coates AS, et al. Personalizing the treatment of women with early breast cancer: highlights of the St Gallen International Expert Consensus on the Primary Therapy of Early Breast Cancer 2013. Ann Oncol. 2013;24(9):2206-23. doi:10.1093/annonc/mdt303
46. Budhathoki S, Hidaka A, Yamaji T, et al. Plasma 25-hydroxyvitamin D concentration and subsequent risk of total and site specific cancers in Japanese population: large case-cohort study within Japan Public Health Center-based Prospective Study cohort. BMJ. 2018;360:k671. doi:10.1136/bmj.k671

47. Deschasaux M, Souberbielle JC, Latino-Martel P, et al. Weight status and alcohol intake modify the association between vitamin D and breast cancer risk. J Nutr. 2016;146(3):576-85. doi:10.3945/jn.115.221481
48. 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):5423-30. doi:10.1158/0008-5472.CAN-16-0353
49. Ordonez-Mena JM, Schottker B, Haug U, et al. Serum 25-hydroxyvitamin D and cancer risk in older adults: results from a large German prospective cohort study. Cancer Epidemiol Biomarkers Prev. 2013;22(5):905-16. doi:10.1158/1055-9965.EPI-12-1332
50. Skaaby T, Husemoen LL, Thuesen BH, 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):1220-9. doi:10.1158/1055-9965.EPI-14-0007
51. Heath AK, Hodge AM, Ebeling PR, et al. Circulating 25-Hydroxyvitamin D concentration and risk of breast, prostate, and colorectal cancers: The Melbourne Collaborative Cohort Study. Cancer Epidemiol Biomarkers Prev. 2019;28(5):900-8. doi:10.1158/1055-9965.EPI-18-1155
52. Acikgoz A, Cimrin D, Ergor G. Effect of serum 25-hydroxyvitamin D level on lung, breast, colorectal and prostate cancers: a nested case-control study. East Mediterr Health J. 2020;26(7):794-802. doi:10.26719/emhj.20.035
53. O'Brien KM, Sandler DP, Taylor JA, Weinberg CR. Serum Vitamin D and risk of breast cancer within five years. Environ Health Perspect. 2017;125(7):077004. doi:10.1289/EHP943
54. McDonnell SL, Baggerly CA, French CB, et al. Breast cancer risk markedly lower with serum 25-hydroxyvitamin D concentrations >/=60 vs <20 ng/ml (150 vs 50 nmol/L): Pooled analysis of two randomized trials and a prospective cohort. PLoS One. 2018;13(6):e0199265. doi:10.1371/journal.pone.0199265
55. Zhu K, Knuiman M, Divitini M, et al. Lower serum 25-hydroxyvitamin D is associated with colorectal and breast cancer, but not overall cancer risk: a 20-year cohort study. Nutr Res. 2019;67:100-7. doi:10.1016/j.nutres.2019.03.010
56. Ordonez-Mena JM, Schottker 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(3):311-23. doi:10.1007/s10654-015-0040-7
57. Peila R, Rohan TE. Association of Prediagnostic Serum Levels of Vitamin D with Risk of Ductal Carcinoma In Situ of the Breast in the UK Biobank Cohort Study. Cancer Epidemiol Biomarkers Prev. 2022;31(7):1499-502. doi:10.1158/1055-9965.EPI-22-0017
58. Zhou L, Chen B, Sheng L, Turner A. The effect of vitamin D supplementation on the risk of breast cancer: a trial sequential meta-analysis. Breast cancer research and treatment. 2020;182(1):1-8. doi:10.1007/s10549-020-05669-4
59. Manson JE, Bassuk SS, Buring JE, Group VR. Principal results of the VITamin D and OmegA-3 TriaL (VITAL) and updated meta-analyses of relevant vitamin D trials. J Steroid Biochem Mol Biol. 2020;198:105522. doi:10.1016/j.jsbmb.2019.105522
60. Keum N, Lee DH, Greenwood DC, Manson JE, Giovannucci E. Vitamin D supplementation and total cancer incidence and mortality: a meta-analysis of

randomized controlled trials. Ann Oncol. 2019;30(5):733-43.
doi:10.1093/annonc/mdz059
61. Manson JE, Cook NR, Lee IM, et al. Vitamin D supplements and prevention of cancer and cardiovascular disease. N Engl J Med. 2019;380(1):33-44.
doi:10.1056/NEJMoa1809944
62. Chandler PD, Chen WY, Ajala ON, et al. Effect of Vitamin D3 Supplements on Development of Advanced Cancer: A Secondary Analysis of the VITAL Randomized Clinical Trial. JAMA Netw Open. 2020;3(11):e2025850.
doi:10.1001/jamanetworkopen.2020.25850
63. McCullough ML, Zoltick ES, Weinstein SJ, et al. Circulating vitamin D and colorectal cancer risk: an international pooling project of 17 cohorts. Journal of the National Cancer Institute. 2019;111(2):158-69. doi:10.1093/jnci/djy087
64. McCullough ML, Weinstein SJ, Freedman DM, et al. Correlates of circulating 25-hydroxyvitamin D: Cohort Consortium Vitamin D Pooling Project of Rarer Cancers. Am J Epidemiol. 2010;172(1):21-35.
65. Hofmann JN, Yu K, Horst RL, Hayes RB, Purdue MP. 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. doi:10.1158/1055-9965.EPI-09-1121
66. Jorde R, Sneve M, Hutchinson M, Emaus N, Figenschau Y, Grimnes G. 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):903-8.
doi:10.1093/aje/kwq005

Table 1 Descriptive characteristics of studies participating in pooled analyses of circulating 25(OH)D and breast cancer

Cohort	Country or continent	Breast cancer cases / controls	25(OH)D data[a]	Median (10%-90%) season-standardized 25(OH)D in controls (nmol/L)[b]	Median (10%-90%) age at blood collection (y)[c]	Years of blood collection[c]	Median (10%-90%) time from blood collection to diagnosis (y)
U.S.							
Beta-Carotene and Retinol Efficacy Trial (CARET)	U.S.	195 / 195	New	41 (19–73)	60 (54–68)	1987-1996	5.5 (1.2–10.3)
CLUE II: Campaign Against Cancer and Heart Disease (CLUE II)	U.S.	447 / 447	New	46 (23–74)	55 (37–71)	1989	9.6 (2.1–17.0)
Cancer Prevention Study-II (CPS-II)	U.S.	413 / 413	Calibrated	60 (33–88)	70 (37–71)	1998-2001	2.3 (0.6–4.9)
Multiethnic Cohort Study (MEC)	U.S.	556 / 556	Calibrated	45 (34–58)	67 (58–79)	1994-2006	3.2 (0.6–6.4)
Nurses' Health Study (NHS)	U.S.	1,368 / 1,368	Calibrated	70 (43–101)	57 (47–66)	1989-1991	7.4 (1.8–15.8)
Nurses' Health Study II (NHSII)	U.S.	415 / 824	Calibrated	67 (42–96)	45 (38–50)	1996-1999	4.6 (0.9–8.4)
New York University Women's Health Study (NYU WHS)	U.S.	893 / 1,642	Calibrated	47 (25–73)	52 (39–63)	1985-1991	11.5 (3.7–17.9)
Prostate, Lung, Colorectal and Ovarian Cancer Screening Trial (PLCO)	U.S.	761 / 975	Calibrated	58 (34–82)	63 (57–70)	1993-2001	3.6 (0.9–7.3)
Women's Health Initiative (WHI)[d]	U.S.	1,003 / 1,003	Calibrated	56 (31–86)	63 (55–73)	1994-2009	3.6 (0.8–6.8)
Women's Health Study (WHS)	U.S.	196 / 196	New	59 (34–92)	49 (46–53)	1992-1995	6.0 (1.4–9.2)
Outside U.S.							
Breakthrough Generations Study (BGS)[e]	U.K.	311 / 311	New	54 (30–82)	58 (48–68)	2004-2009	1.7 (0.4–3.6)
Etude Epidémiologique auprès de femmes de l'Education Nationale (E3N)	France	582 / 1,074	Calibrated	31 (20–47)	56 (49–66)	1994-1999	3.8 (0.8–7.1)
European Investigation into Cancer and Nutrition (EPIC)	Europe	1,229 / 1,229	Calibrated	43 (23–68)	50 (39–61)	1992-2002	4.5 (1.1–8.6)
JANUS Serum Bank (JANUS)	Norway	388 / 388	Calibrated	71 (42–103)	41 (39–52)	1973-2002	12.4 (3.6–20.3)
Malmö Diet and Cancer Study (MDCS)	Sweden	590 / 590	Calibrated	60 (47–77)	56 (48–66)	1991-1996	7.2 (1.8–12.1)
Northern Sweden Health and Disease Study (NSHDS)	Sweden	691 / 1,296	Calibrated	47 (29–70)	56 (50–67)	1995-2006	6.1 (0.8–11.9)
Hormones and Diet in the Etiology of Breast Cancer (ORDET)	Italy	446 / 446	New	39 (24–66)	49 (39–62)	1987-1992	9.1 (2.0–15.3)
Total		10,484 / 12,953		51 (27–83)[f]	57 (42–68)	1973-2009	5.1 (1.1–13.4)

25(OH)D: 25-hydroxyvitamin D.

[a] New studies: Circulating 25(OH)D concentrations were measured in cases and controls using a direct, competitive chemiluminescence immunoassay at Heartland Assays (Ames, IA) between November 2011 and April 2013. Calibrated studies: In these studies, circulating 25(OH)D concentrations had been previously measured and needed to be calibrated. For each study, blood samples from 29 controls selected uniformly across deciles of the 25(OH)D distribution in the study were re-assayed during 2011–2013 with the same assay and in the same laboratory used for the new studies.
[b] Includes newly measured and calibrated circulating 25(OH)D data. To convert circulating 25(OH)D from nmol/L to ng/mL, divide values by 2.496.
[c] For cases and controls combined.
[d] WHI randomized trial of calcium and vitamin D.
[e] Now called Generations Study.
[f] For all cases combined, median (interdecile range) circulating 25(OH)D was 52 (26-84) nmol/L.

Table 2 RRs of breast cancer by consortium-wide 25(OH)D deciles and continuous 25(OH)D

Deciles of 25(OH)D (nmol/L)[a]	Cases / controls	Median 25(OH)D in controls (nmol/L)	RR (95% CI)		
			Model 1[b]	Model 2[c]	Model 3[d]
<27	976 / 1,295	22.6	1.04 (0.91-1.20)	1.02 (0.88-1.17)	1.03 (0.89-1.19)
27 - <34	978 / 1,296	31.0	1.02 (0.89-1.16)	1.01 (0.88-1.15)	1.01 (0.89-1.16)
34 - <40	1,065 / 1,295	37.5	1.10 (0.97-1.24)	1.09 (0.96-1.23)	1.11 (0.98-1.26)
40 - <46	1,027 / 1,295	43.0	1.02 (0.90-1.14)	1.01 (0.89-1.14)	1.01 (0.89-1.14)
46 - <51	1,036 / 1,295	48.5	1.0 (ref)[e]	1.0 (ref)[e]	1.0 (ref)[e]
51 - <57	1,078 / 1,296	54.0	1.04 (0.92-1.17)	1.04 (0.93-1.18)	1.04 (0.91-1.18)
57 - <63	1,105 / 1,295	60.1	1.05 (0.93-1.18)	1.08 (0.96-1.22)	1.11 (0.98-1.26)
63 - <71	1,096 / 1,295	66.9	1.00 (0.88-1.13)	1.02 (0.90-1.16)	1.04 (0.92-1.19)
71 - <83	1,044 / 1,295	76.5	0.96 (0.85-1.10)	1.00 (0.88-1.15)	1.00 (0.87-1.15)
≥83	1,079 / 1,296	93.2	1.01 (0.88-1.16)	1.05 (0.92-1.21)	1.05 (0.91-1.21)
P-trend[f]			0.20	0.83	0.64
P-heterogeneity for decile 1[g]			0.90	0.89	0.67
P-heterogeneity for decile 10[g]			0.55	0.63	0.52
Continuous 25(OH)D per 25 nmol/L			0.98 (0.94-1.02)	1.00 (0.96-1.03)	0.99 (0.95-1.03)
P-heterogeneity[g]			0.35	0.68	0.63

25(OH)D: 25-hydroxyvitamin D; BMI: body mass index; CI: confidence interval; ref: referent; RR: relative risk.

[a] Consortium-wide decile cut-points were based on the 25(OH)D distribution in controls from all studies combined.

[b] Model 1 conditioned on study-specific matching factors, which included age at blood collection and usually date of blood collection. The matching factors for each study are provided in **Supplementary Table S1**.

[c] Model 2 added to model 1 BMI (using a variable combining BMI and menopausal status at blood collection) and physical activity.

[d] Model 3 further adjusted for established breast cancer risk factors: race, family history of breast cancer, age at menarche, parity/age at first birth, oral contraceptive use, menopausal hormone therapy, and alcohol intake.

[e] Decile 5 was chosen as the referent category since it permitted us to estimate RR at low and high concentrations and included cases and controls from each of the participating studies.

[f] *P*-trend was calculated with a Wald test that assigned the median 25(OH)D in controls to each decile.

[g] *P*-heterogeneity in the results across studies was evaluated using the Q statistic.

Legends for Figures 1-5

Fig. 1 Median circulating 25(OH)D concentrations among controls for the 12 calibrated studies, before and after calibration, and for the five newly measured studies. For the calibrated studies, median 25(OH)D before calibration is indicated by red circles and after calibration by blue circles. For the newly measured studies, median 25(OH)D is indicated by blue squares. All 25(OH)D concentrations are season-standardized. The relative change after calibration is +10% for CPS-II, -47% for E3N, -18% for EPIC, +16% for JANUS, -31% for MDCS, -41% for MEC, +14% for NHS, +9% for NHSII, -11% for NSHDS, -14% for NYU WHS, -9% for PLCO, and +11% for WHI. Horizontal green lines indicate IOM-based clinical guidance for circulating 25(OH)D: deficiency at <30 nmol/L, insufficiency at 30-<50 nmol/L, sufficiency at 50-<75 mmol/L, beyond sufficiency at ≥75 nmol/L. The full names of the 17 cohorts are given in **Table 1**. 25(OH)D: 25-hydroxyvitamin D; IOM: Institute of Medicine.

Fig 2 RRs of breast cancer by 25(OH)D categories based on IOM clinical guidance for vitamin D. RRs (solid circles) of breast cancer and 95% CIs (vertical lines) are plotted on a log scale for 11 categories of circulating 25(OH)D. These 25(OH)D categories expand those established by the IOM. The categories <20 (3.1% of controls) and 20-<30 (10.4%) nmol/L are considered deficient; 30-<40 (15.7%) and 40-<50 (18.5%) nmol/L are considered insufficient; 50-<62.5 (21.2%) (the referent) and 62.5-<75 (14.9%) nmol/L are considered sufficient; and 75-<87.5 (9.0%), 87.5-<100 (3.9%), 100-<112.5 (1.8%), 112.5-<125 (0.9%), and ≥125 nmol/L (0.7%) are considered beyond sufficient. The RR and 95% CI for each category are plotted at the median concentration of 25(OH)D among controls in that category. The model used, model 3, conditioned on study-specific matching factors, which included age at blood collection and usually date of blood collection, and adjusted for BMI (using a variable combining BMI and menopausal status at blood collection), physical activity, race, family history of breast cancer, age at menarche, parity and age at first birth, oral contraceptive use, menopausal hormone therapy, and alcohol intake. The *P*-trend across categories was calculated with a Wald test that assigned the median 25(OH)D in controls to each category. 25(OH)D: 25-hydroxyvitamin D; BMI: body mass index; CI: confidence interval; IOM: Institute of Medicine; ref: referent; RR: relative risk.

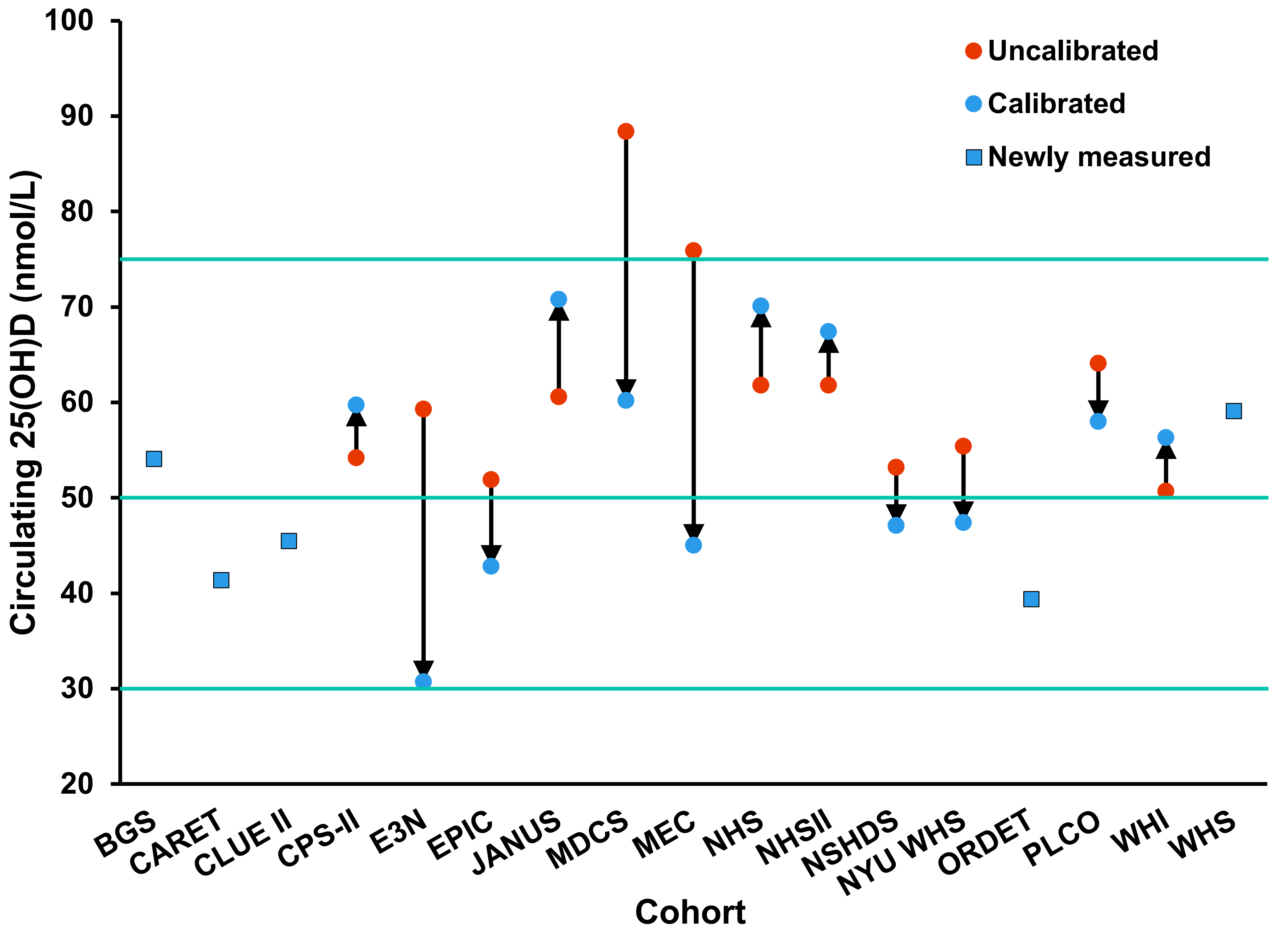
Fig 3 Cohort-specific and pooled RRs of breast cancer per 25 nmol/L increase in circulating 25(OH)D. Cohort-specific RRs of breast cancer per 25 nmol/L increase in 25(OH)D are plotted on a log scale and indicated by solid squares with the size proportional to the cohort's contribution to the pooled RR. Horizontal lines indicate 95% CIs. The pooled RR is indicated by a solid diamond. A RR of 1.0, marked by the vertical line, indicates no association. The model used, model 3, conditioned on study-specific matching factors, which included age at blood collection and usually date of blood collection, and adjusted for BMI (using a variable combining BMI and menopausal status at blood collection), physical activity, race, family history of breast cancer, age at menarche, parity and age at first birth, oral contraceptive use, menopausal hormone therapy, and alcohol intake. The *P*-heterogeneity *in* the results across cohorts was calculated using the Q statistic. The full names of the 17 cohorts are given in **Table 1**.

25(OH)D: 25-hydroxyvitamin D; BMI: body mass index; CI: confidence interval; ref: referent; RR: relative risk.

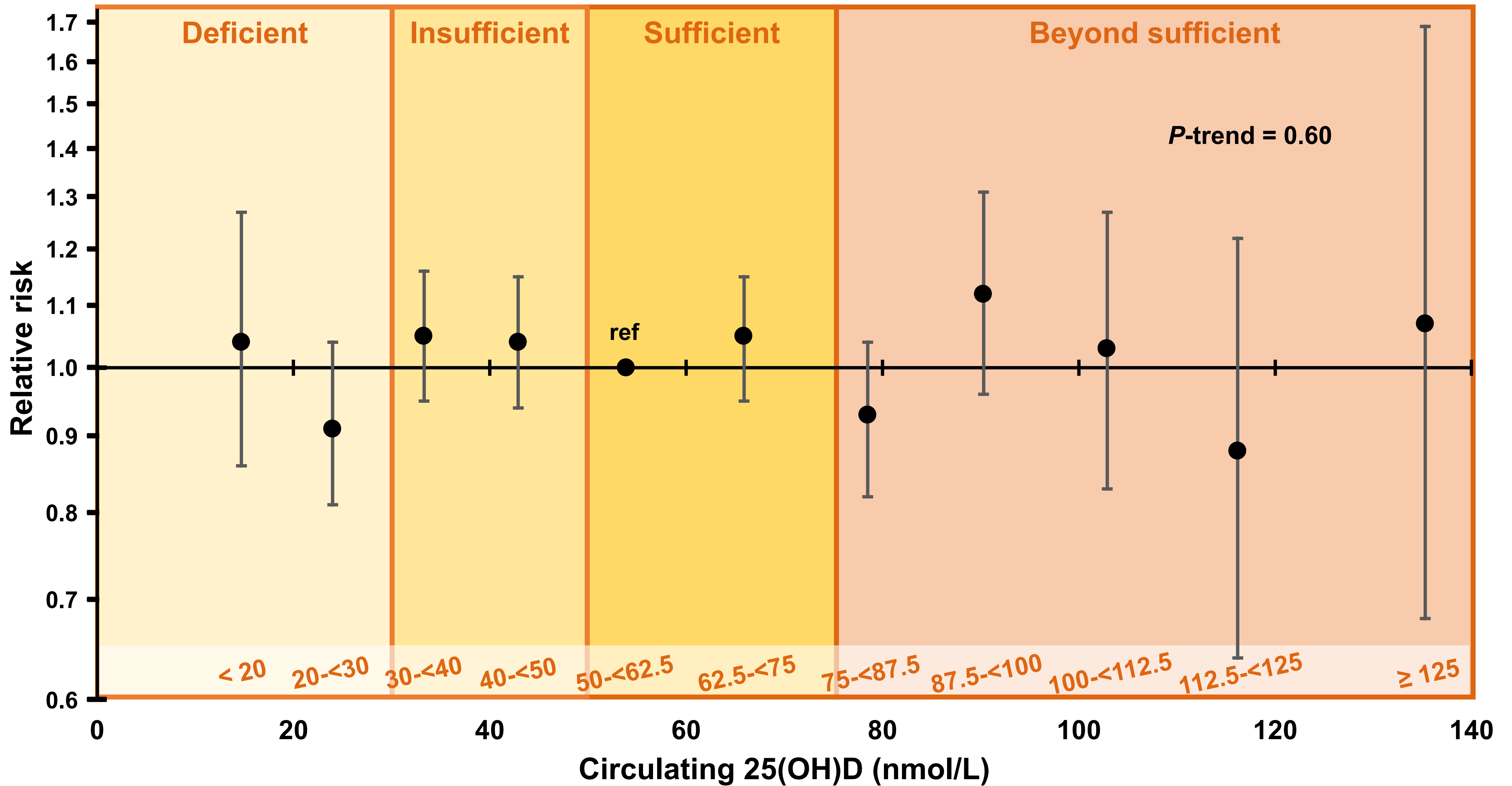
Fig. 4 RRs of breast cancer per 25 nmol/L increase in circulating 25(OH)D in population subgroups and by tumor characteristics. RRs (solid squares) of breast cancer and 95% CIs (horizontal lines) per 25 nmol/L increase in 25(OH)D are plotted on a log scale. A RR of 1.0, marked by the vertical line, indicates no association. Differences in RR by participant characteristics, such as breast cancer risk factors, latitude, and season of blood collection, were examined using a two-stage approach and unconditional logistic regression. Differences in RR by study characteristics, such as geographic region and calibration (newly measured vs. calibrated 25(OH)D data), were examined using a two-stage approach and conditional logistic regression. Differences in associations by participant and study characteristics were tested for statistical significance using meta-regression. Heterogeneity by time elapsed from blood collection to diagnosis, age at diagnosis, and tumor characteristics were assessed in an aggregated dataset with conditional analyses, and tested for statistical significance with a contrast test. Conditional models were conditioned on study-specific matching factors, which included age at blood collection and usually date of blood collection; while unconditional models included as covariates age at blood collection, elapsed time from blood collection to assay, and study center (EPIC and MEC only). In addition, all models were further adjusted for BMI (using a variable combining BMI and menopausal status at blood collection), physical activity, race, family history of breast cancer, age at menarche, parity and age at first birth, oral contraceptive use, menopausal hormone therapy, and alcohol intake. Analyses by use of menopausal hormone therapy were restricted to postmenopausal and perimenopausal women. For analyses stratified by season of blood collection, winter was defined as weeks 2-14; summer as weeks 28-40; and circulating 25(OH)D was not season-standardized. In the analyses by molecular subtype, ER+ and/or PR+/HER2- tumors are considered luminal A-like, ER+ and/or PR+/HER2+ tumors are considered luminal B-like, ER-/PR-/HER2+ tumors are considered HER2-enriched, and ER-/PR-/HER2- tumors are considered triple negative. Additional information on how we defined and harmonized breast tumor characteristics is presented in **Supplementary Methods**. 25(OH)D: 25-hydroxyvitamin D; BMI: body mass index; CI: confidence interval; ER: estrogen receptor; HER2: human epidermal growth factor receptor 2; PR: progesterone receptor; ref: referent; RR: relative risk; +; positive; −: negative.

Fig 5 RRs of breast cancer by season-specific deciles of circulating 25(OH)D. RRs (solid circles) of breast cancer and 95% CIs (vertical lines) are plotted on a log scale by winter-specific and summer-specific, consortium-wide deciles of circulating 25(OH)D. Results for the participants who donated blood in the winter (defined as weeks 2-14; 2,741 cases) are shown in blue; results for those who donated in the summer (defined as weeks 28-40; 2,491 cases) are shown in orange. Participants donating blood in the spring and fall were excluded from these analyses. 25(OH)D concentrations were not season-standardized. Decile cutpoints, based on the 25(OH)D distributions in controls, were for winter <22, 22-<28, 28-<34, 34-<39, 39-<45, 45-<50, 50-<56, 56-<63, 63-<75, and ≥75 nmol/L and for summer <35, 35-<43, 43-<50, 50-<55, 55-<61, 61-<67, 67-<72,

72-<81, 81-<94, and ≥94 nmol/L. The RR and 95% CI for each decile was plotted at the median 25(OH)D concentration among the controls in that decile. Unconditional logistic regression models in an aggregated dataset were adjusted for study, age at blood collection, elapsed time from blood collection to assay, study center (EPIC and MEC only), BMI (using a variable combining BMI and menopausal status at blood collection), physical activity, race, family history of breast cancer, age at menarche, parity and age at first birth, oral contraceptive use, menopausal hormone therapy, and alcohol intake. The *P*-trend across categories was based on a Wald test that assigned the median 25(OH)D in controls to each category. 25(OH)D: 25-hydroxyvitamin D; BMI: body mass index; CI: confidence interval; N: number; ref: referent.



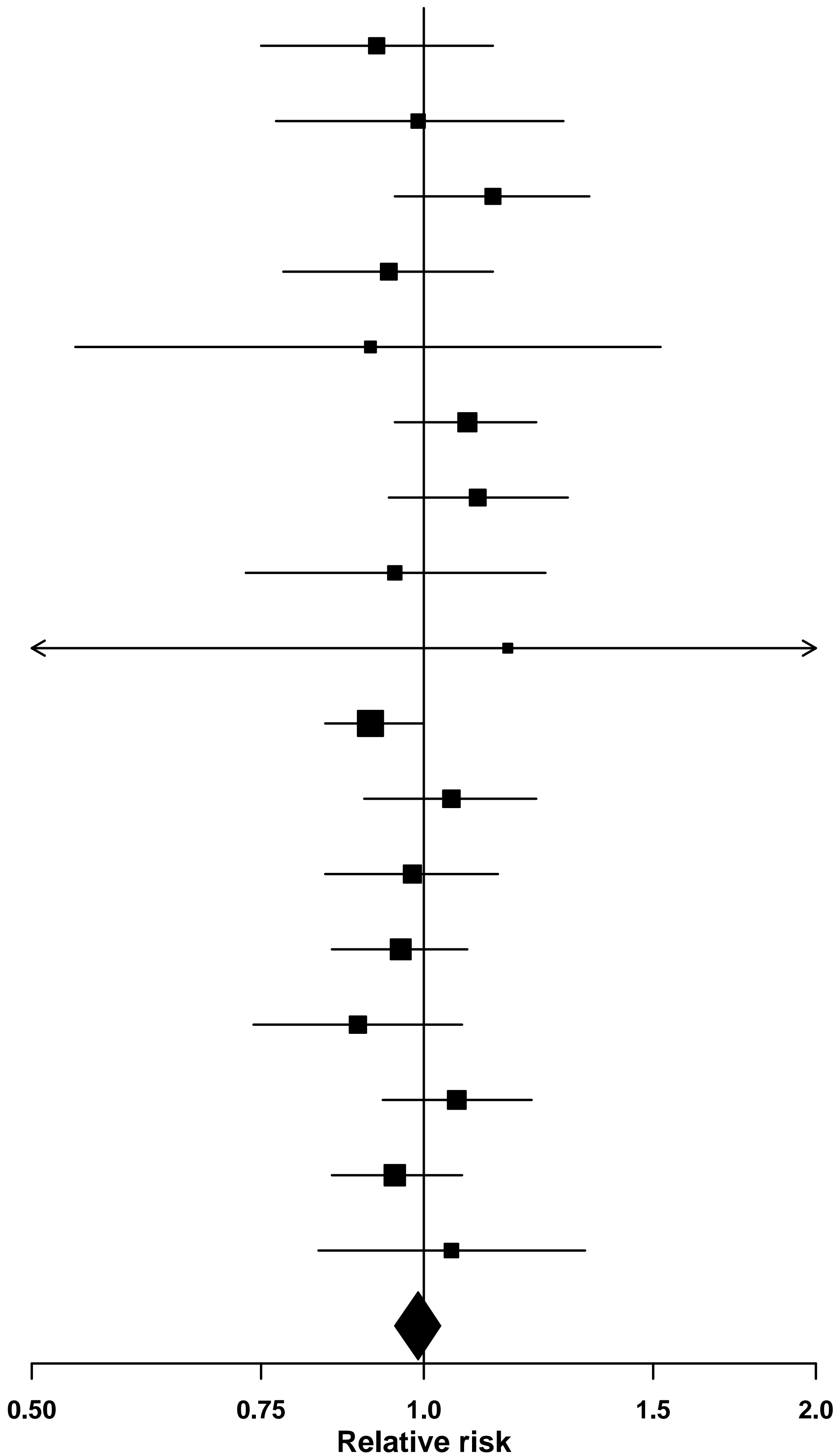
Uncalibrated
Calibrated
Newly measured
Circulating 25(OH)D (nmol/L)
100
90
80
70
60
50
40
30
20
BGS
CARET
CLUE II
CPS-II
E3N
EPIC
JANUS
MDCS
MEC
NHS
NHSII
NSHDS
NYU WHS
ORDET
PLCO
WHI
WHS
Cohort

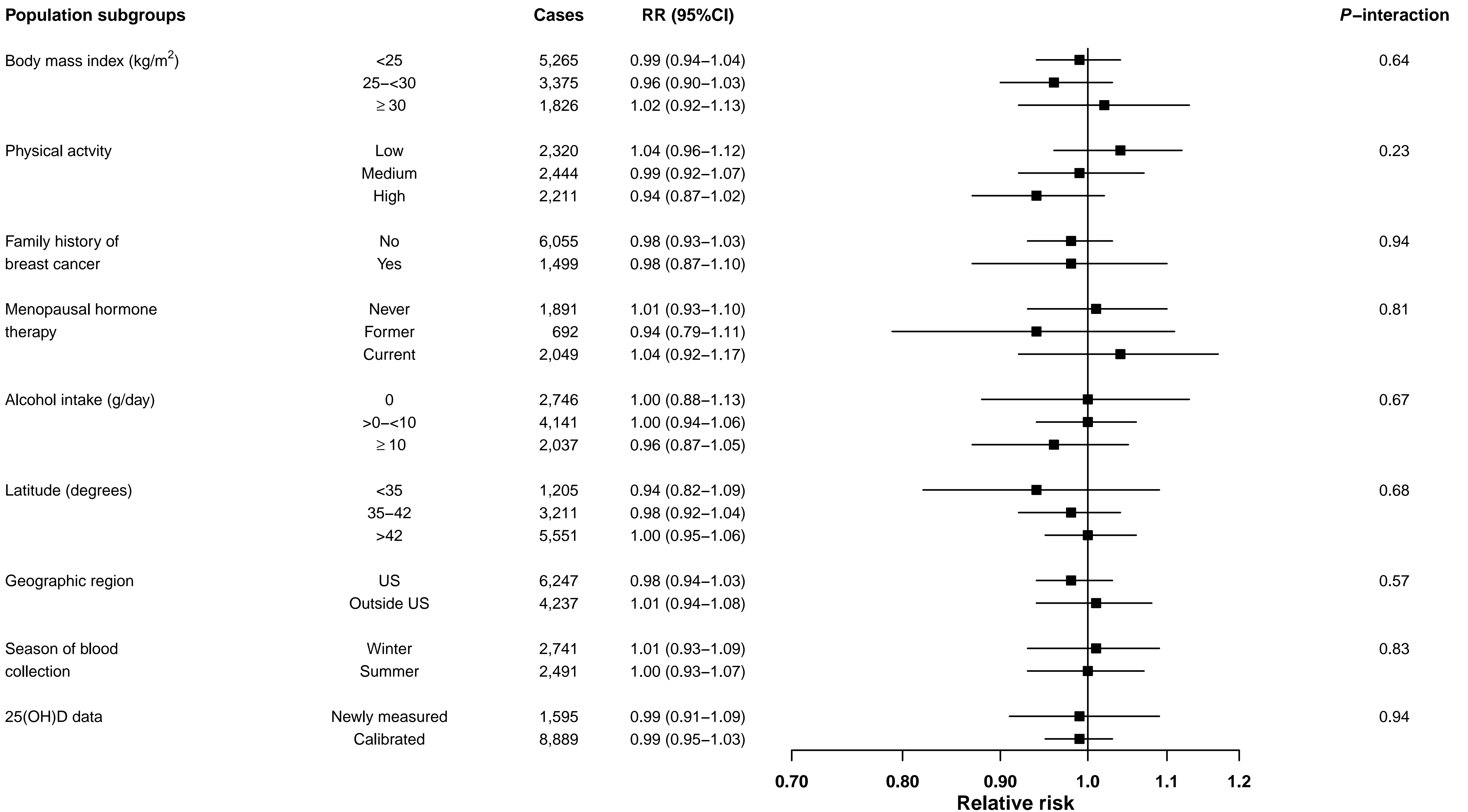


Deficient
Insufficient
Sufficient
Beyond sufficient
P-trend = 0.60
ref
Relative risk
1.7
1.6
1.5
1.4
1.3
1.2
1.1
1.0
0.9
0.8
0.7
0.6
< 20
20-<30
30-<40
40-<50
50-<62.5
62.5-<75
75-<87.5
87.5-<100
100-<112.5
112.5-<125
≥ 125
0
20
40
60
80
100
120
140
Circulating 25(OH)D (nmol/L)

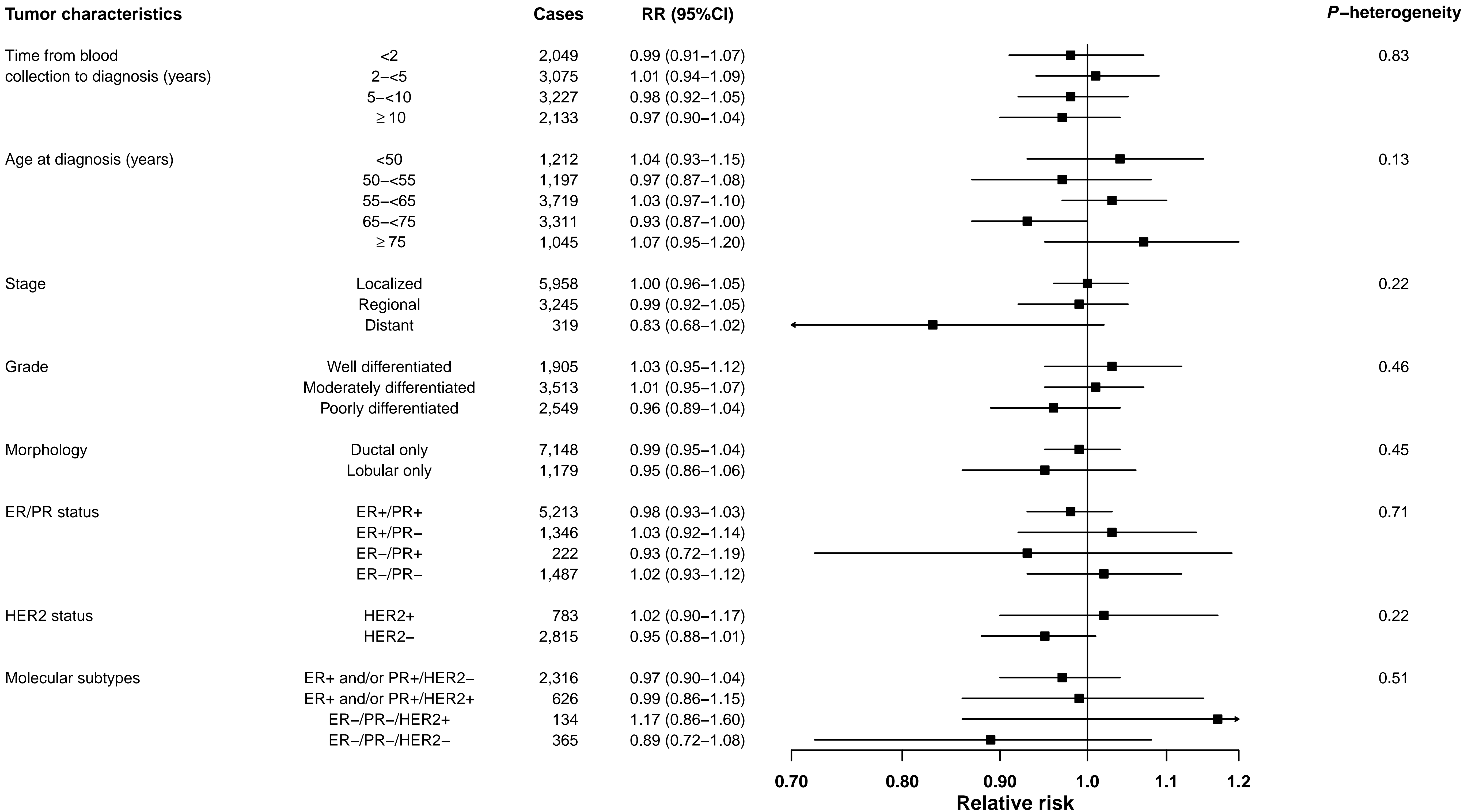
Cohort	RR (95%CI)
BGS	0.92 (0.75–1.13)
CARET	0.99 (0.77–1.28)
CLUE II	1.13 (0.95–1.34)
CPS–II	0.94 (0.78–1.13)
E3N	0.91 (0.54–1.52)
EPIC	1.08 (0.95–1.22)
JANUS	1.10 (0.94–1.29)
MDCS	0.95 (0.73–1.24)
MEC	1.16 (0.38–3.49)
NHS	0.91 (0.84–1.00)
NHSII	1.05 (0.90–1.22)
NSHDS	0.98 (0.84–1.14)
NYU WHS	0.96 (0.85–1.08)
ORDET	0.89 (0.74–1.07)
PLCO	1.06 (0.93–1.21)
WHI	0.95 (0.85–1.07)
WHS	1.05 (0.83–1.33)
Pooled	**0.99 (0.95–1.03)**

***P*–heterogeneity = 0.63**

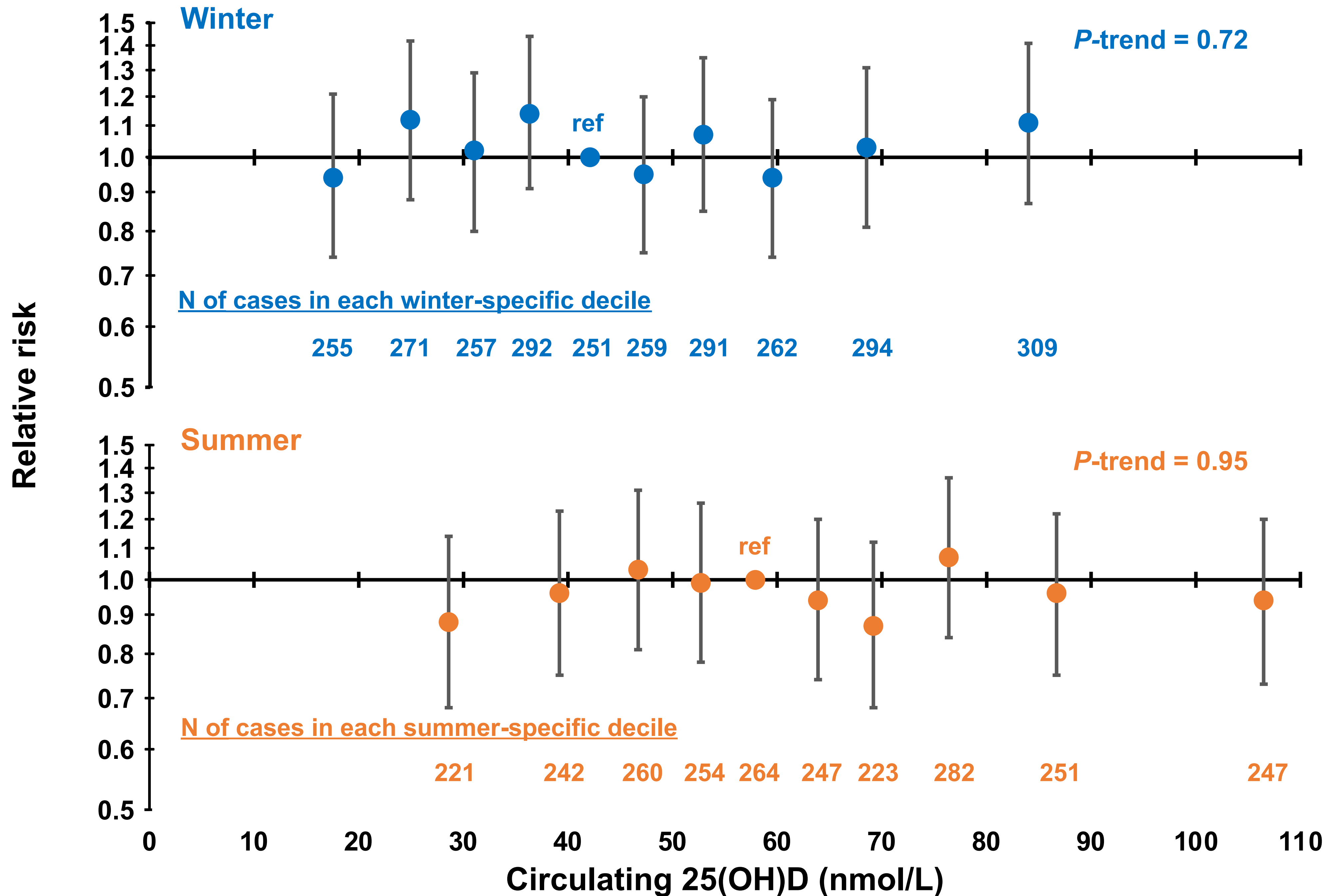




Population subgroups
Cases
RR (95%CI)
P−interaction
Body mass index (kg/m2)
<25
5,265
0.99 (0.94−1.04)
0.64
25−<30
3,375
0.96 (0.90−1.03)
≥ 30
1,826
1.02 (0.92−1.13)
Physical actvity
Low
2,320
1.04 (0.96−1.12)
0.23
Medium
2,444
0.99 (0.92−1.07)
High
2,211
0.94 (0.87−1.02)
Family history of
breast cancer
No
6,055
0.98 (0.93−1.03)
0.94
Yes
1,499
0.98 (0.87−1.10)
Menopausal hormone
therapy
Never
1,891
1.01 (0.93−1.10)
0.81
Former
692
0.94 (0.79−1.11)
Current
2,049
1.04 (0.92−1.17)
Alcohol intake (g/day)
0
2,746
1.00 (0.88−1.13)
0.67
>0−<10
4,141
1.00 (0.94−1.06)
≥ 10
2,037
0.96 (0.87−1.05)
Latitude (degrees)
<35
1,205
0.94 (0.82−1.09)
0.68
35−42
3,211
0.98 (0.92−1.04)
>42
5,551
1.00 (0.95−1.06)
Geographic region
US
6,247
0.98 (0.94−1.03)
0.57
Outside US
4,237
1.01 (0.94−1.08)
Season of blood
collection
Winter
2,741
1.01 (0.93−1.09)
0.83
Summer
2,491
1.00 (0.93−1.07)
25(OH)D data
Newly measured
1,595
0.99 (0.91−1.09)
0.94
Calibrated
8,889
0.99 (0.95−1.03)
0.70
0.80
0.90
1.0
1.1
1.2
Relative risk



Tumor characteristics
Cases
RR (95%CI)
P−heterogeneity
Time from blood collection to diagnosis (years)
<2
2,049
0.99 (0.91−1.07)
0.83
2−<5
3,075
1.01 (0.94−1.09)
5−<10
3,227
0.98 (0.92−1.05)
≥ 10
2,133
0.97 (0.90−1.04)
Age at diagnosis (years)
<50
1,212
1.04 (0.93−1.15)
0.13
50−<55
1,197
0.97 (0.87−1.08)
55−<65
3,719
1.03 (0.97−1.10)
65−<75
3,311
0.93 (0.87−1.00)
≥ 75
1,045
1.07 (0.95−1.20)
Stage
Localized
5,958
1.00 (0.96−1.05)
0.22
Regional
3,245
0.99 (0.92−1.05)
Distant
319
0.83 (0.68−1.02)
Grade
Well differentiated
1,905
1.03 (0.95−1.12)
0.46
Moderately differentiated
3,513
1.01 (0.95−1.07)
Poorly differentiated
2,549
0.96 (0.89−1.04)
Morphology
Ductal only
7,148
0.99 (0.95−1.04)
0.45
Lobular only
1,179
0.95 (0.86−1.06)
ER/PR status
ER+/PR+
5,213
0.98 (0.93−1.03)
0.71
ER+/PR−
1,346
1.03 (0.92−1.14)
ER−/PR+
222
0.93 (0.72−1.19)
ER−/PR−
1,487
1.02 (0.93−1.12)
HER2 status
HER2+
783
1.02 (0.90−1.17)
0.22
HER2−
2,815
0.95 (0.88−1.01)
Molecular subtypes
ER+ and/or PR+/HER2−
2,316
0.97 (0.90−1.04)
0.51
ER+ and/or PR+/HER2+
626
0.99 (0.86−1.15)
ER−/PR−/HER2+
134
1.17 (0.86−1.60)
ER−/PR−/HER2−
365
0.89 (0.72−1.08)
0.70
0.80
0.90
1.0
1.1
1.2
Relative risk



Winter
P-trend = 0.72
ref
N of cases in each winter-specific decile
255
271
257
292
251
259
291
262
294
309
Summer
P-trend = 0.95
ref
N of cases in each summer-specific decile
221
242
260
254
264
247
223
282
251
247
Relative risk
1.5
1.4
1.3
1.2
1.1
1.0
0.9
0.8
0.7
0.6
0.5
0
10
20
30
40
50
60
70
80
90
100
110
Circulating 25(OH)D (nmol/L)

Supplementary Methods

Measurement of circulating 25(OH)D

For each of the 12 calibrated studies (**Table 1**), blood samples from 29 controls, selected uniformly across deciles of the 25(OH)D distribution in study controls, were re-assayed at Heartland Assays (Ames, IA). Three controls were selected at random from each decile but one (either the fifth or sixth decile) of the 25(OH)D distribution. Then for each calibrated study, a M-estimation robust linear regression model(35) was used to convert its 25(OH)D data to the concentrations expected had the Heartland Assays chemiluminescence immunoassay been used.(36, 37) The extra variation due to fitting the calibration models was taken into account using previously published methods.(36, 37)

Laboratory personnel were blinded to case-control status. Two types of blinded quality controls were randomly inserted into each assay batch by the individual cohorts: 1) National Institute of Standards and Technology (NIST) Standard Reference Materials with known 25(OH)D concentrations and 2) cohort-specific quality controls chosen by the cohort. Coefficients of variation (CVs) for the chemiluminescence immunoassay, including within- and between-batch variability across both newly assayed and calibrated studies, were 13.4%, 7.8%, and 7.8%, respectively, for NIST reference materials at 17.7, 32.3, and 49.8 nmol/L concentrations. CVs, including within- and between-batch variability, ranged from 5.3% to 14.7% for the study-specific quality controls inserted into the newly assayed studies.

For 10 of the 12 studies to be recalibrated, blood samples were re-assayed with the same direct, competitive chemiluminescence immunoassay kits used for the five newly measured studies (DiaSorin LIAISON 25 OH Vitamin D TOTAL; DiaSorin, Stillwater, MN).(33, 34) However, for two of the studies we needed two calibration steps. For E3N and MEC, a direct, competitive radioimmunoassay (DiaSorin 25 OH Vitamin D ^{125}I RIA; DiaSorin, Stillwater, MN)(33, 34) had to be utilized to re-assay their blood samples since these cohorts joined the breast cancer project after the original chemiluminescence immunoassay kits had expired and the new version of the chemiluminescence immunoassay was not optimized for heparin-plasma samples, the matrix utilized by MEC. Therefore, the original 25(OH)D measurements for E3N and MEC were first calibrated to the radioimmunoassay by re-assaying blood samples from 29 controls in each study. Then the radioimmunoassay was calibrated to the chemiluminescence immunoassay with 60 serum samples from PLCO controls. These 60 samples, previously measured for 25(OH)D with our chemiluminescence immunoassay kits at Heartland Assays as part of our consortium's study of vitamin D and colorectal cancer,(67) were re-assayed with the radioimmunoassay concurrently with the E3N and MEC samples. Study-specific M-estimation robust linear regression models(35) were used to calibrate the original 25(OH)D data from E3N and MEC to the radioimmunoassay and to calibrate those results to the chemiluminescence immunoassay.(36, 37)

For the radioimmunoassay, CVs, including within- and between-batch variability, were 6.8%, 8.1%, and 6.6% for NIST reference materials at 17.7, 32.3, and 49.8 nmol/L concentrations, respectively.

Since two calibrations were necessary to convert the original 25(OH)D data from E3N and MEC to the chemiluminescence assay, we reran the analyses for consortium-wide deciles, IOM-based categories, and continuous 25(OH)D after excluding the

participants from E3N (5.6% of cases) and MEC (5.3% of cases). Results were essentially unchanged (data not shown).

Breast tumor characteristics

All cases were primary invasive breast cancer (ICD 9 170.0-174.9; ICD 10 C50.0 C50.9). Each cohort provided information on the tumor characteristics of its individual cases; these data were harmonized at the BBC3 coordinating center.

Stage: Cohorts classified breast cancer cases by stage using two different systems. TNM staging(68, 69) was used by 14 cohorts, and Surveillance, Epidemiology, and End Results Program (SEER) staging(70) by three cohorts. We harmonized the stage information into three categories: localized, regional, and distant. The algorithm below outlines how this was done for TNM data. For SEER data, we combined regional due to direct extension, regional due to lymph node involvement, and regional due to both direct extension and lymph node involvement. Stage was available for 9,522 cases (91%).

Stage	**TNM**			**Assigned extent of disease**
	Tumor size	**Number of nodes**	**Metastasis**	
1	T1: ≤2 cm	0	No or blank	Localized
		>0		Regional
2	T2: >2 – 5 cm	0	No or blank	Localized
		>0		Regional
3	T3: >5 cm	0	No or blank	Localized
		>0		Regional
	T4: any size and direct extension to chest wall or skin	≥0		Regional
4	Any T or blank	≥0	Yes	Distant

Grade: Individual cohorts provided information on tumor grade based on two systems. Eight cohorts classified tumors as well differentiated, moderately differentiated, poorly differentiated, or undifferentiated. For our analyses, we combined tumors that were poorly differentiated and undifferentiated. Seven cohorts used a breast cancer-specific grading system, specifically the Nottingham Histologic Score System(71), also termed the Elston-Ellis modification of the Scarff-Bloom-Richardson grading system,(72) which classifies tumors as well differentiated, moderately differentiated, or poorly differentiated. Information on grade was not available for two cohorts. Grade was available for 7,967 cases (76%).

Morphology: Fourteen cohorts provided ICDO morphology codes(73) for the cases. We identified the tumors reported to have only ductal carcinoma cells (8500) or only lobular carcinoma cells (8520). For the two cohorts that classified tumors as ductal, lobular, or mixed, we included only the ductal and lobular tumors. Information on morphology was not provided by one cohort. Of the 9,787 breast cancer cases with morphology data (93%

of all cases), 7,148 (73%) were ductal only and 1,179 (12%) were lobular only.

Estrogen receptor (ER) and progesterone receptor (PR) status: All but one cohort provided information on ER and PR status. Hormone receptor status was classified as either positive, negative, or missing or as positive, borderline, negative, or missing. Following the American Society of Clinical Oncology/College of American Pathologists guidelines,(74) which state that ER and PR assay results should be considered positive if at least 1% of the invasive tumor nuclei in the sample test positive, we combined borderline with positive readings. Of the 8,788 breast cancer cases with ER data (84% of all cases), 6,949 (79%) were ER-positive (ER+). Of the 8,295 cases with PR data (79% of all cases), 5,444 (66%) were PR-positive (PR+).

Human epidermal growth factor receptor 2 (HER2) status: Eleven of the 17 cohorts provided information on HER2 status. Tumors were classified as positive, borderline, negative, or missing. We excluded the borderline tumors as these readings could be positive or negative based on further testing.(75) Of the 3,689 breast cancer cases with HER2 data (35% of all cases), 783 (21%) were HER2-positive (HER2+), and 2,815 (76%) were HER2-negative (HER2-).

***Molecular subtypes*:** Based on hormone receptor data from 11 cohorts, we classified tumors as luminal A-like (ER+ and/or PR+/HER2−), luminal B-like (ER+ and/or PR+/HER2+), HER2-enriched (ER−/PR−/HER2+), or triple negative (ER−/PR−/HER2−).(45)

Statistical analysis

Three models were used to control for confounding. Model 1 conditioned on study-specific matching factors, which included age at blood collection and usually date of blood collection (**Supplementary Table S1).** Alternatively, in unconditional models, age at blood collection, elapsed time from blood collection to assay, and for EPIC and MEC, study center were included as covariates in Model 1. Model 2 added to model 1 BMI, as a combination BMI and menopausal status variable (<25 kg/m^2 and premenopausal, 25 – <30 kg/m^2 and premenopausal, ≥30 kg/m^2 and premenopausal, <25 kg/m^2 and peri/postmenopausal, 25 –<30 kg/m^2 and peri/postmenopausal, ≥30 kg/m^2 and peri/postmenopausal) and physical activity (study-specific tertiles of metabolic equivalents, if this level of detail was available, or categorized into low, moderate, high). A variable combining BMI categories with menopausal status at date of blood collection was used because breast cancer risk is inversely associated with premenopausal BMI and positively associated with postmenopausal BMI.(44) Model 3, the model presented unless otherwise noted, was further adjusted for established breast cancer risk factors: race (white, black, Asian, other; only for studies that did not match on race), family history of breast cancer (no, yes), age at menarche (continuous), parity and age at first birth (nulliparous, 1-2 children and <25 y, 1-2 children and ≥25 y, ≥3 children and <25 y, ≥3 children and ≥25 y), oral contraceptive use (never, former, current; or never, ever; or not current, current; based on the information collected), menopausal hormone therapy (never, former, current; or never, ever; or not current, current; based on the information collected), and alcohol intake (continuous).

All covariates included a missing category, except race, BMI, and menopausal status. Women for whom race was not specified were categorized as white. Women missing BMI were assigned the median BMI for study participants from their cohort. For women missing menopausal status, we used 51 y, the mean age of menopause,(76) to estimate menopausal status. Women <51 y at blood collection were coded as premenopausal; women ≥51 y at blood collection were coded as peri/postmenopausal.

Although we used a two-stage approach to estimate RRs in most analyses, for selected analyses we aggregated the data from all studies into a single dataset. In general, we utilized an aggregated dataset when one or more studies contributed very few cases and/or controls to a specific stratum. Aggregated datasets were used to compare risk between extreme deciles; to assess heterogeneity by tumor characteristics, time elapsed from blood collection to diagnosis, age at diagnosis, and race; to explore the importance of season of blood collection with analyses restricted to women who donated blood in the winter or summer; and to test the linearity of the 25(OH)D-breast cancer association with restricted cubic spline analyses. When unconditional models were used in an aggregated dataset, study was included as a covariate.

Differences in the circulating 25(OH)D-breast cancer associations by participant characteristics, such as breast cancer risk factors, latitude, and season of blood collection, were examined using a two-stage approach and unconditional logistic regression. Differences by study characteristics, such as geographic region and calibration (newly measured vs. calibrated 25(OH)D data), were also examined using a two-stage approach but with conditional logistic regression. Differences in associations by participant and study characteristics were tested for statistical significance using meta-regression.(77) In all two-stage analyses, individual studies with fewer than 25 cases in a stratum were excluded from the stratum. Because of the small numbers of blacks and Asians, heterogeneity by race was evaluated in an aggregated dataset by including race-25(OH)D interaction terms in unconditional analyses. A Wald test was used to test for statistical significance. Heterogeneity by time elapsed from blood collection to diagnosis, age at diagnosis, and tumor characteristics was assessed in an aggregated dataset with conditional analyses because of small numbers of some breast cancer subtypes, and tested for statistical significance with a contrast test.(78)

For analyses by season of blood donation, winter and summer were empirically defined as the 13 weeks centered on the week with the lowest and highest, respectively, predicted circulating 25(OH)D concentration in controls. First sine-cosine functions were used to model 25(OH)D concentrations as a function of week of blood donation for the controls in each study.(37, 38) Then the study-specific predicted 25(OH)D concentrations for each week were weighted by study size and averaged. Using this approach, winter was defined as weeks 2-14, and summer as weeks 28-40.

We initially examined whether the season of blood donation (winter or summer) modified breast cancer risk using 25(OH)D as a continuous variable. To explore further the influence of season, we stratified circulating 25(OH)D into winter-specific and summer-specific, consortium-wide deciles, based on the distributions among consortium controls, and examined the 25(OH)D-breast cancer relationship among those women who donated blood in the winter (2,741 cases) and those women who donated in the summer (2,491 cases). Unconditional logistic regression was conducted in an aggregated dataset to accommodate the small numbers of participants from some studies who donated blood in

a specific season. For season-specific analyses, 25(OH)D concentrations were not season-standardized.

References

67. Weinstein SJ, Purdue MP, Smith-Warner SA, et al. Serum 25-hydroxyvitamin D, vitamin D binding protein and risk of colorectal cancer in the Prostate, Lung, Colorectal and Ovarian Cancer Screening Trial. Int J Cancer. 2015;136(6):E654-64. doi:10.1002/ijc.29157

68. Amin MB, American Joint Committee on Cancer., American Cancer Society. AJCC cancer staging manual. Eight edition / editor-in-chief, Mahul B. Amin, MD, FCAP ; editors, Stephen B. Edge, MD, FACS and 16 others ; Donna M. Gress, RHIT, CTR - Technical editor ; Laura R. Meyer, CAPM - Managing editor. ed. Chicago IL: American Joint Committee on Cancer, Springer; 2017.

69. Edge SB, American Joint Committee on Cancer. AJCC cancer staging manual. 7th ed. New York: Springer; 2010.

70. National Cancer Institute. SEER Summary Staging Manual 2000. Surveillance, Epidemiology, and End Results Program. 2012. https://seer.cancer.gov/tools/ssm/ssm2000/. Accessed July 2020.

71. Elston CW, Ellis IO. Pathological prognostic factors in breast cancer. I. The value of histological grade in breast cancer: experience from a large study with long-term follow-up. Histopathology. 1991;19(5):403-10. doi:10.1111/j.1365-2559.1991.tb00229.x

72. Bloom HJ, Richardson WW. Histological grading and prognosis in breast cancer; a study of 1409 cases of which 359 have been followed for 15 years. Br J Cancer. 1957;11(3):359-77. doi:10.1038/bjc.1957.43

73. National Cancer Institute. Morphology & Grade. SEER Training Modules. https://training.seer.cancer.gov/breast/abstract-code-stage/morphology.html. Accessed July 2020.

74. Hammond ME, Hayes DF, Dowsett M, et al. American Society of Clinical Oncology/College Of American Pathologists guideline recommendations for immunohistochemical testing of estrogen and progesterone receptors in breast cancer. J Clin Oncol. 2010;28(16):2784-95. doi:10.1200/JCO.2009.25.6529

75. Wolff AC, Hammond MEH, Allison KH, et al. Human Epidermal Growth Factor Receptor 2 Testing in Breast Cancer: American Society of Clinical Oncology/College of American Pathologists Clinical Practice Guideline Focused Update. J Clin Oncol. 2018;36(20):2105-22. doi:10.1200/JCO.2018.77.8738

76. Research on the menopause in the 1990s. Report of a WHO Scientific Group. World Health Organ Tech Rep Ser. 1996;866:1-107.

77. Stram DO. Meta-analysis of published data using a linear mixed-effects model. Biometrics. 1996;52:536-44.

78. Wang M, Spiegelman D, Kuchiba A, et al. Statistical methods for studying disease subtype heterogeneity. Stat Med. 2016;35(5):782-800. doi:10.1002/sim.6793

Supplementary Tables S1 – S3

Supplementary Table 1 Additional characteristics of studies participating in pooled analyses of circulating 25(OH)D and breast cancer

Cohort[a]	Case-control matching factors	Vitamin D assay
BGS	Individual matching 1:1 Matched on: • age at blood collection • year of blood collection • ethnicity • days between blood collection and processing	Direct competitive chemiluminescence immunoassay by Heartland Assays, Inc. (Ames, IA, USA) using the DiaSorin LIAISON platform (DiaSorin, Inc., Stillwater, MN, USA).
CARET	Individual matching 1:1 Matched on: • age at blood collection • date of blood collection • race/ethnicity	Direct competitive chemiluminescence immunoassay by Heartland Assays, Inc. (Ames, IA, USA) using the DiaSorin LIAISON platform (DiaSorin, Inc., Stillwater, MN, USA).
CLUE II	Individual matching 1:1 Matched on: • birth date • date of blood collection • race • time since last menstrual period prior to blood collection • hour of blood collection • Washington County residency	Direct competitive chemiluminescence immunoassay by Heartland Assays, Inc. (Ames, IA, USA) using the DiaSorin LIAISON platform (DiaSorin, Inc., Stillwater, MN, USA).
CPS-II	Individual matching 1:1 Matched on: • birth date • date of blood collection • race/ethnicity	Direct competitive chemiluminescence immunoassay by Heartland Assays, Inc. (Ames, IA, USA) using the DiaSorin LIAISON platform (DiaSorin, Stillwater, MN, USA).

	• menopausal status at blood collection (all postmenopausal)	
E3N	Individual matching 1:2 Matched on: • age at blood collection • year of blood collection • menopausal status at blood collection • age at menopause • study center	Chemiluminescence immunoassay by the biochemistry laboratory of Bichat Hospital (Paris, France) using the Elecsys Analyzer (Roche Diagnostic, Florham Park, NJ, USA).
EPIC	Individual matching 1:1 Matched on: • age at blood collection • menopausal status at blood collection • phase of the menstrual cycle at blood collection • exogenous hormone use at blood collection • time of day at blood collection • fasting status at blood collection • study center	Enzyme immunoassay by the immunoassay laboratory in the Cancer Epidemiology Division of the German Cancer Research Center (DFKZ; Heidelberg, Germany) using the OCTEIA platform (Tyne and Wear, UK).
JANUS	Individual matching 1:1 Matched on: • age at blood collection • date of blood collection	Competitive radioimmunoassay by the hormone laboratory of Oslo University Hospital Aker (Oslo, Norway) using the DiaSorin LIAISON platform (DiaSorin, Stillwater, MN, USA).
MDCS	Individual matching 1:1 Matched on: • age at blood collection • date of blood collection • menopausal status at blood collection	High-pressure liquid chromatography in the Department of Clinical Chemistry of Malmö University Hospital (Malmö, Sweden) using a Chromsystems instrument (Chromsystems Instruments & Chemicals GmbH, Gräfelfing, Germany).

<table>
<tr><td>MEC</td><td>Individual matching 1:1
Matched on:
birth year
date of blood draw
race/ethnicity
menopausal status at blood collection (all postmenopausal)
time of day at blood draw
hours fasting prior to blood draw
menopausal hormone therapy use at blood collection
geographic location</td><td>Isotope dilution liquid chromatography-mass spectrometry in the laboratory of Dr. Adrian Franke at the University of Hawaii Cancer Center (Mānoa, HI, USA) using an Orbitrap mass spectrometer (ThermoFisher Scientific, Waltham, MA, USA).</td></tr>
<tr><td>NHS</td><td>Individual matching 1:1
Matched on:
age at blood collection
month of blood collection
menopausal status at blood collection and diagnosis
menopausal hormone therapy use at blood collection and diagnosis
time of day at blood collection
fasting status at blood collection</td><td>Radioimmunoassay in the laboratory of Dr. Bruce Hollis at the University of South Carolina (Columbia, SC, USA), protein-binding assay in the laboratory of Dr. Michael Holick at the Boston University School of Medicine (Boston, MA, USA), and a direct competitive chemiluminescence immunoassay by Heartland Assays (Ames, IA, USA) using the DiaSorin LIAISON platform (DiaSorin, Inc., Stillwater, MN, USA).</td></tr>
<tr><td>NHSII</td><td>Individual matching 1:2
Matched on:
age at blood collection
month/year of blood collection
race/ethnicity
menopausal status at blood collection (all premenopausal)
menopausal status at diagnosis
luteal day of menstrual cycle</td><td>Radioimmunoassay after acetonitrile extraction by Heartland Assays, Inc. (Ames, IA, USA).</td></tr>
</table>

	• time of day at blood collection • fasting status at blood collection	
NSHDS	Individual matching 1:2 Matched on: • age at first blood collection • date of first blood collection • number and dates of subsequent blood collections	Direct competitive chemiluminescence immunoassay by Heartland Assays, Inc. (Ames, IA, USA) using the DiaSorin LIAISON platform (DiaSorin, Inc., Stillwater, MN, USA).
NYU WHS	Individual matching 1:2 Matched on: • age at first blood collection • date of first blood collection • race/ethnicity • menopausal status at blood collection • number and dates of subsequent blood collections	Direct competitive chemiluminescence immunoassay by Heartland Assays, Inc. (Ames, IA, USA) using the DiaSorin LIAISON platform (DiaSorin, Inc., Stillwater, MN, USA).
ORDET	Individual matching 1:1 Matched on: • age at blood collection • date of blood collection • menopausal status at blood collection	Direct competitive chemiluminescence immunoassay by Heartland Assays, Inc. (Ames, IA, USA) using the DiaSorin LIAISON platform (DiaSorin, Inc., Stillwater, MN, USA).
PLCO	Frequency matching 1:1 Matched on: • age at blood collection • date of blood collection	Radioimmunoassay in the laboratory of Dr. Bruce Hollis at the University of South Carolina (Columbia, SC, USA).
WHI	Individual matching 1:1 Matched on: • age at enrollment • date of blood collection • race/ethnicity • menopausal status at blood collection (all postmenopausal)	Direct competitive chemiluminescence immunoassay in the laboratory of Dr. Bruce Hollis at the University of South Carolina (Columbia, SC, USA) using the DiaSorin LIAISON platform (DiaSorin, Inc., Stillwater, MN, USA).

	• latitude of the clinical center • WHI trial arm	
WHS	Individual matching 1:1 Matched on: • age at blood collection • date of blood collection • race • menopausal status at blood collection • time of day at blood collection • fasting status	Direct competitive chemiluminescence immunoassay by Heartland Assays, Inc. (Ames, IA, USA) using the DiaSorin LIAISON platform (DiaSorin, Inc., Stillwater, MN, USA).

25(OH)D: 25-hydroxyvitamin D.

[a] The full names of the 17 cohorts are given in **Table 1**.

Supplementary Table 2 RRs of breast cancer by 25(OH)D categories based on IOM clinical guidance[a]

25(OH)D categories (nmol/L)	Cases / controls	RR (95% CI)		
		Model 1[b]	Model 2[c]	Model 3[d]
<20	329 / 400	1.11 (0.92-1.33)	1.05 (0.87-1.27)	1.04 (0.86-1.27)
20 - <30	960 / 1,348	0.95 (0.84-1.07)	0.92 (0.81-1.04)	0.91 (0.81-1.04)
30 - <40	1,623 / 2,028	1.06 (0.96-1.16)	1.03 (0.94-1.14)	1.05 (0.95-1.16)
40 - <50	1,992 / 2,401	1.07 (0.96-1.18)	1.04 (0.94-1.15)	1.04 (0.94-1.15)
50 - <62.5	2,218 / 2,741	1.0 (ref)	1.0 (ref)	1.0 (ref)
62.5 - <75	1,638 / 1,924	1.03 (0.94-1.12)	1.04 (0.95-1.14)	1.05 (0.95-1.15)
75 - <87.5	890 / 1,171	0.92 (0.83-1.03)	0.94 (0.84-1.06)	0.93 (0.82-1.04)
87.5 - <100	461 / 503	1.10 (0.94-1.28)	1.13 (0.97-1.31)	1.12 (0.96-1.31)
100 - <112.5	208 / 232	1.06 (0.87-1.31)	1.08 (0.88-1.33)	1.03 (0.83-1.27)
112 - <125	84 / 115	0.85 (0.62-1.16)	0.88 (0.64-1.21)	0.88 (0.64-1.22)
≥125	81 / 90	1.05 (0.68-1.62)	1.09 (0.70-1.69)	1.07 (0.68-1.69)
P-trend[e]		0.33	0.88	0.60
P-heterogeneity for <20 nmol/L[f]		0.73	0.66	0.57
P-heterogeneity for ≥125 nmol/L[f]		0.27	0.28	0.25

25(OH)D: 25-hydroxyvitamin D; BMI: body mass index; CI: confidence interval; IOM: Institute of Medicine; ref: referent; RR: relative risk.

[a] We used an expanded version of the IOM recommendations for vitamin D, which are based on bone health. Circulating 25(OH)D <30 nmol/L is considered deficient, 30-<50 nmol/L is considered insufficient, 50-<75 nmol/L is considered sufficient, and ≥75 nmol/L is considered beyond sufficient

[b] Model 1 conditioned on study-specific matching factors, which included age at blood collection and usually date of blood collection. The matching factors for each study are provided in **Supplementary Table S1**.

[c] Model 2 added to model 1 BMI (using a variable combining BMI and menopausal status at blood collection) and physical activity.

[d] Model 3 further adjusted for established breast cancer risk factors: race, family history of breast cancer, age at menarche, parity/age at first birth, oral contraceptive use, menopausal hormone therapy, and alcohol intake.

[e] *P*-trend was calculated with a Wald test that assigned the median 25(OH)D in controls to each category.

[f] *P*-heterogeneity in the results across studies was evaluated using the Q statistic.

Supplementary Table 3 Comparison of two-stage and aggregated models for estimating pooled RRs of breast cancer by circulating 25(OH)D

25(OH)D measure	RR (95% CI)[a]		
	Two-stage random effects models	Two-stage fixed effects models	Aggregated models[b]
Consortium-wide deciles (nmol/L)[c]			
<27	1.03 (0.89-1.19)	1.03 (0.89-1.19)	1.02 (0.90-1.16)
27 - <34	1.01 (0.89-1.16)	1.01 (0.89-1.16)	1.01 (0.90-1.15)
34 - <40	1.11 (0.98-1.26)	1.11 (0.98-1.26)	1.06 (0.94-1.20)
40 - <46	1.01 (0.89-1.14)	1.01 (0.89-1.14)	1.00 (0.89-1.12)
46 - <51	1.0 (ref)	1.0 (ref)	1.0 (ref)
51 - <57	1.04 (0.91-1.18)	1.04 (0.92-1.18)	1.03 (0.91-1.15)
57 - <63	1.11 (0.98-1.26)	1.11 (0.98-1.26)	1.06 (0.95-1.20)
63 - <71	1.04 (0.92-1.19)	1.04 (0.92-1.19)	1.03 (0.92-1.17)
71 - <83	1.00 (0.87-1.15)	1.00 (0.87-1.15)	0.99 (0.87-1.11)
≥83	1.05 (0.91-1.21)	1.05 (0.91-1.21)	1.00 (0.88-1.13)
P-trend[d]	0.64	0.64	0.63
Categories based on IOM guidance (nmol/L)[e]			
<20	1.04 (0.86-1.27)	1.04 (0.86-1.27)	1.08 (0.91-1.28)
20 - <30	0.91 (0.81-1.04)	0.91 (0.81-1.04)	0.97 (0.86-1.08)
30 - <40	1.05 (0.95-1.16)	1.05 (0.95-1.16)	1.05 (0.96-1.15)
40 - <50	1.04 (0.94-1.15)	1.04 (0.95-1.14)	1.05 (0.96-1.14)
50 - <62.5	1.0 (ref)	1.0 (ref)	1.0 (ref)
62.5 - <75	1.05 (0.95-1.15)	1.05 (0.95-1.15)	1.04 (0.95-1.14)
75 - <87.5	0.93 (0.82-1.04)	0.93 (0.82-1.04)	0.92 (0.82-1.02)
87.5 - <100	1.12 (0.96-1.31)	1.12 (0.96-1.31)	1.11 (0.96-1.29)
100 - <112.5	1.03 (0.83-1.27)	1.03 (0.83-1.27)	1.06 (0.87-1.30)
112 - <125	0.88 (0.64-1.22)	0.88 (0.64-1.22)	0.85 (0.63-1.15)
≥125	1.07 (0.68-1.69)	1.00 (0.70-1.45)	1.05 (0.77-1.45)
P-trend[d]	0.60	0.60	0.63
Continuous			
per 25 nmol/L	0.99 (0.95-1.03)	0.99 (0.95-1.03)	0.99 (0.95-1.03)

25(OH)D: 25-hydroxyvitamin D; BMI: body mass index; CI: confidence interval; IOM: Institute of Medicine; ref: referent; RR: relative risk.

[a] Model 3. Conditioned on study-specific matching factors, which included age at blood collection and usually date of blood collection, and adjusted for BMI (using a variable combining BMI and menopausal status at blood collection), physical activity, race, family history of breast cancer, age at menarche, parity/age at first birth, oral contraceptive use, menopausal hormone therapy, and alcohol intake. The matching factors for each study are provided in **Supplementary Table S1**.

[b] For these aggregated models, the data from all studies were combined into a single dataset and analyzed together, while stratifying for study.

[c] Consortium-wide decile cut-points were based on the 25(OH)D distribution in controls from all studies combined.

[d] *P*-trend was calculated with a Wald test that assigned the median 25(OH)D in controls to each category.

[e] Expanded version of the IOM recommendations for vitamin D, which are based on bone health. Circulating 25(OH)D <30 nmol/L is considered deficient, 30-<50 nmol/L is considered insufficient, 50-<75 nmol/L is considered sufficient, and ≥75 nmol/L is considered beyond sufficient.