

<https://doi.org/10.1038/s41523-024-00702-x>

Plasma C-peptide mammographic features and risk of breast cancer



Shadi Azam¹✉, Cheng Peng², Bernard A. Rosner², Marcus D. Goncalves³, Erica Phillips³, Heather Eliassen^{2,4}, John Heine⁵, Susan E. Hankinson⁶ & Rulla M. Tamimi¹

Our study in the Nurses' Health Study (NHS) and NHS2, a nested case-control study with 1260 cases and 2221 controls, investigated the association between C-peptide levels, mammographic density (MD) parameters, V (a measure of gray scale variation), and breast cancer (BC) risk. We also examined how C-peptide and BC risk vary across quartiles of mammographic features. Linear and logistic regressions were used to study the associations between C-peptide and MD parameters, and breast cancer. C-peptide was inversely associated with percent MD and positively with non-dense area, but no associations were found with dense area and V measure. C-peptide was associated with an increased risk of invasive BC risk (top vs. bottom quartile, odds ratio = 1.46, 95% CI: 1.12–1.91). No multiplicative interactions were found between C-peptide, MD parameters, and BC risk. Our results suggest a positive association between C-peptide and BC risk, and MD parameters do not seem to modify this association.

C-peptide (connecting peptide) connects alpha and beta chains of proinsulin. It is released by the beta cells within the islets of Langerhans in the endocrine pancreas during the conversion of proinsulin into insulin and C-peptide¹. C-peptide has a longer half-life in plasma than insulin, the degradation rate of it in the body is lower than of insulin and therefore C-peptide levels may be more stable and useful in epidemiological studies to reflect the degree of insulin resistance and be less affected by β cell dysfunction of a given individual². There are several lines of evidence to suggest that metabolic hormones such as C-peptide and insulin may have mitotic effects on both normal and neoplastic breast epithelial cells, stimulate the proliferation of human breast cell, and enhances tumor cell migration and invasion capacity³.

To date, there have been several prospective studies investigating the potential association between circulating insulin or C-peptide levels and the risk of breast cancer. Results have been inconsistent, suggesting an increased risk with elevated insulin levels or C-peptide^{4–8} or no association^{9–12}. Insulin increases the bioavailability of insulin-like growth factor I (IGF-I), a hormone that regulates the effects of growth hormone (GH) and promotes the normal growth of bones and tissues. In both in vitro and in vivo studies, insulin and IGF-1 can synergistically stimulate proliferation in mammary cells¹³. In our prior nested case-control study in the Nurses' Health Study (NHS) and NHSII, high levels of circulating C-peptide were associated with

an approximately 50% increased risk of invasive breast cancer [top vs. bottom quartile, adjusted OR = 1.5, 95% confidence interval (CI), 1.1–2.0]. In stratified analyses, the results were similar for breast cancers diagnosed before and after menopause⁴. Additionally, findings from the Women's Health Initiative Observational Study limited to only postmenopausal women revealed that high fasting insulin levels (lowest vs. highest quartile) were statistically significantly associated with a 46% increased risk of breast cancer comparing the highest to the lowest quartile of insulin level¹⁴. A meta-analysis including 11 prospective studies reported a 26% increase in the risk of breast cancer incidence with elevated insulin or C-peptide blood levels (pooled relative risk RR = 1.26, 95% CI 1.06–1.48)¹⁵. Further, high levels of C-peptide were associated with breast hyperplasia (overgrowth of the lobular and ductal cells) in postmenopausal women¹¹. The physiological mechanism involved in breast hyperplasia and changes in breast tissue associated with high C-peptide levels may be explained by upregulators of ovarian steroid secretion and other growth factors such as IGF¹⁶. Mammographic density is a well-established risk factor for breast cancer^{17,18}. By definition, the dense part of the breast consists of epithelial tissue and stroma and appears bright on a mammogram, whereas fat tissue appears dark. Previous studies showed women of the same age and body mass index (BMI), with very dense breasts, have a greater risk of breast cancer by 4–6 times as compared with women with less than 5% dense breast tissue^{19,20}. In

¹Department of Population Health Sciences, Weill Cornell Medicine, New York, NY, USA. ²Channing Division of Network Medicine, Brigham and Women's Hospital, Boston, MA, USA. ³Department of Medicine, Weill Cornell Medicine, New York, NY, USA. ⁴Departments of Nutrition and Epidemiology, Harvard T.H. Chan School of Public Health, Boston, MA, USA. ⁵Department of Cancer Epidemiology, H. Lee Moffitt Cancer Center & Research Institute, Tampa, FL, USA.

⁶Department of Biostatistics and Epidemiology, School of Public Health and Health Sciences, University of Massachusetts Amherst, Amherst, MA, USA. ✉e-mail: sha4015@med.cornell.edu



addition to mammographic density parameters, several studies have demonstrated that mammographic texture features are associated with breast cancer risk independent of mammographic density^{21–24}. A texture summary measure called “V” captures gray-scale variation in mammograms and was shown to be a stronger predictor of breast cancer risk than percent mammographic density in three Mayo Clinic cohort studies²⁵.

To date, only three studies have evaluated the association between plasma C-peptide levels and mammographic density^{4,26,27} and the results of all these studies indicated no association after adjustment for adiposity. However, these studies including our own earlier report, had smaller sample sizes than our current study (ranging from $n = 95–1785$)^{4,26,27}, and did not stratify analyses by menopausal status^{26,27}. In the current study, we aimed to investigate the association between plasma C-peptide concentration, mammographic density parameters, mammographic texture feature, and overall breast cancer risk in two well-characterized prospective case-control studies nested within the NHS/NHSII cohorts. We specifically assessed associations by breast cancer subtype and stratified analyses by menopausal status. Furthermore, we assessed whether mammographic density parameters and/or V measure modify the association between C-peptide and the risk of breast cancer.

Methods

Study population

The study included women from the Nurses’ Health Study (NHS) and NHSII. NHS was initiated in 1976 when 121,700 registered nurses in the United States who were aged 30–55 years returned an initial mailed questionnaire with detailed information on their medical and reproductive history. The cohort was limited to married women due to the sensitivity of questions about contraceptive use at that time²⁸. The information on BMI, reproductive history, age at menopause, menopausal hormone therapy (MHT) use, and any diagnoses of cancer or other disease were updated biennially using questionnaires, and the response rate at each questionnaire cycle has been greater than 90%. Between 1989 and 1990, 32,826 women ages 43–69 years in NHS provided blood samples²⁹.

NHSII started in 1989 and included 116,429 female registered nurses aged 25–42 at baseline. Similar to NHS, the NHSII participants completed a biennial questionnaire with approximately a 90% response rate. A subset of women ($N = 29,611$) who were cancer-free between the ages of 32–54 years, provided a blood sample between 1996–1999³⁰.

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 as well as those of participating cancer registries as required. The voluntary return of the questionnaires and blood samples were considered to imply consent. This study complied with all relevant ethical regulations including the Declaration of Helsinki.

Case-control selection

We conducted a nested case-control study within the NHS blood sub-cohort, consisting of women diagnosed with invasive breast cancer between blood collection matched to one or two controls. In NHS, 1616 cases were diagnosed between blood collection and the 2004 follow-up cycle and matched to 2438 control subjects on age (± 2 years), menopausal status at blood draw and at diagnosis, month (± 1 month) and time of the day (± 2 h), and fasting status at blood collection at blood collection. NHSII cases ($N = 450$) were diagnosed between blood collection and the 2005 follow-up cycle and matched to controls 1:2 ($N = 893$) based on the following criteria: age (± 2 years), race, menopausal status at blood draw and at diagnosis, as well as luteal day (± 1 day for timed samples), time of the day (± 2 h), month (± 1 month) and fasting status at blood collection.

Measurement of plasma C-peptide concentration

Detailed information on the collection, processing, and storage of NHS and NHSII blood specimens is reported in detail elsewhere³¹. In brief, blood

samples from both cohorts were shipped overnight on ice and separated into plasma, erythrocyte, and leukocyte fractions upon arrival. Plasma samples are stored in liquid nitrogen freezers (< -130 °C). Ten batches of C-peptide were assayed from plasma samples using enzyme-linked immunosorbent assay (Diagnostic Systems Laboratory, Webster, TX) in the laboratory of Dr. Michael Pollak (McGill University, Montréal, Québec, Canada). Matched cases/control sets were assayed and handled together. The laboratory personnel were blinded to the case/control status of all samples³². The percent coefficient of variation (CV) for C-peptide measurement ranged from 5.3% to 9.0% (mean = 6.7%) over 10 assay batches. Plasma C-peptide concentrations were categorized into batch-specific quartiles due to inter-batch variability. The quartile cut points for each batch are presented in Supplementary Table 1.

Mammogram collection and processing

We collected pre-diagnostic, film-screening mammograms from eligible women in the nested case-control study through the 2004 (NHS), and 2005 (NHSII) cycles. We aimed to collect the mammogram closest in date to the blood draw date. The median time difference between blood collection and mammogram dates was -7 months (i.e., blood collection preceded mammogram by 7 months). Because mammogram collections in NHS only went through the 2004 follow-up cycle, we successfully obtained mammograms from approximately 65% of women with C-peptide measurements in the nested case-control study with C-Peptide measurements. Previous results also show that breast cancer risk factors are similar for women we are able to collect mammograms compared to those for whom we are unable to collect a mammogram^{33,34}. Among women with C-peptide measured, 1330 cases and 2223 controls from the 2004 (NHS)/2005 (NHSII) cycles had mammographic density parameters measured. We excluded 70 cases, whose mammography dates were after, or within 1-month preceding breast diagnosis. The study participant flowchart with reasons for exclusion are given in Fig. 1.

From the mammograms collected, the cranio-caudal views of both breasts were digitized at 261 $\mu\text{m}/\text{pixel}$ with a Lumysis 85 laser film scanner (Lumysis, Sunnyvale, CA) or a VIDAR CAD PRO Advantage scanner (VIDAR Systems Corporation; Herndon, VA) (using a comparable resolution of 150 dots per inch and 12-bit depth inch and 12-bit depth). The correlation between percent mammographic density from the two scanners was very high ($r = 0.88$)³⁵.

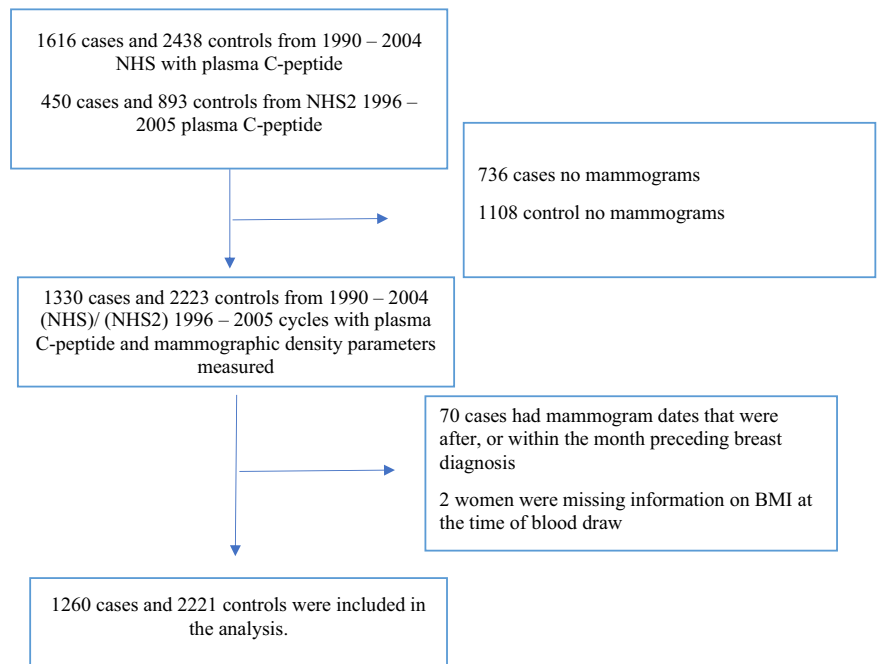
Measurement of mammographic density and texture feature

We used computer-assisted thresholding to measure dense area, non-dense area (dense area subtracted from the total area), and percent mammographic density calculated as dense area divided by the total breast area (Cumulus software, University of Toronto, Toronto, ON, CA)^{36,37}. Analyses of these measurements were based on values averaged across left and right breasts. Mammographic density of the right and left breast is highly correlated¹⁹. A detailed description of mammographic density measurement in the NHS and NHSII participants is reported in an earlier publication³⁵. The mammogram reader was blinded to the subject’s case/control status and biomarker data. Two observers read the mammograms from NHS participants in two batches. For NHSII, a single observer read mammograms in three batches. There was high reproducibility within each batch³⁶. The summary measure of texture feature “V” is an automated measure that captures gray-scale variation on a mammogram. The algorithm and the method have been described previously²⁵.

Covariates

We considered established breast cancer risk factors as potential confounders in the present analysis³⁸. The following information on covariates was obtained from the questionnaire closest to the time of blood collection: age (years), MHT use (postmenopausal and never used MHT, postmenopausal and formerly used MHT, and postmenopausal and currently using MHT), age at menopause (years), age at menarche (years), BMI (kg/m^2), waist-to-hip-ratio, age at first birth and parity (nulliparous, 1 or 2

Fig. 1 | Flow chart describing the exclusion criteria for women participated in the NHS1 and NHS2.



children and <25 years old, 1 or 2 children and 25+, 3+ children and <5 years old, 3+ children and 25+ years old), alcohol consumption (grams/day), physical activity (metabolic equivalent of task [MET]-h/week), history of benign breast diseases (no, yes), family history of breast cancer (no, yes), and diabetes status (no, yes). Women reporting either no menstrual periods within the previous 12 months with natural menopause, bilateral oophorectomy, or hysterectomy with 1 or more ovaries retained and were 54 years or older if a smoker or 56 years or older if a nonsmoker were considered postmenopausal^{39,40}. Percent mammographic density is inversely associated with BMI⁴¹ and positively associated with breast cancer risk⁴². Therefore, it is necessary to adjust for BMI in any study of percent mammographic density and breast cancer risk.

Information on breast cancer

Incident breast cancer cases were identified by self-report on biennial questionnaires and confirmed through a review of medical records. Greater than 99% of reported breast cancer cases were confirmed upon review. Information regarding tumor characteristics (i.e., hormone receptor status) was collected through a central review of stained breast cancer tissue microarrays or by pathology reports when tissue was not available.

Statistical analysis

Descriptive statistics for established determinants of mammographic density and/or breast cancer risk factors among breast cancer controls according to quartiles of plasma C-peptide concentration shows in Table 1. The distributions of dense area, non-dense area, and percent mammographic density were skewed. Square-root transformation yielded normal distributions for each of these parameters. We used linear regression models with 95% confidence intervals to estimate the association of plasma C-peptide and square root transformed mammographic density parameters and V measure among control subjects with and without adjustment for potential confounders. Conditional logistic regression with adjustment for matching factors and potential confounders was used to model associations between the quartile of plasma C-peptide and breast cancer risk as well as tumor hormone receptor status. Tests for heterogeneity were performed with Cochran's Q test⁴³. All *p* value tests for trend were conducted using the Wald test, where the medians of the quartiles were modeled continuously. To examine whether mammographic density parameters and the V

measure modify the associations between C-peptide concentrations and breast cancer risk on the multiplicative scale, we conducted likelihood ratio tests by comparing two generalized linear models with and without the interaction term. Women in the lowest quartile of mammographic density parameters and lowest quartile of C-peptide concentration were used as the reference. Lastly, we performed sensitivity analyses restricted to those without diabetes diagnosis.

Results

Baseline characteristics

Baseline characteristics for the 2221 NHS/NHSII breast cancer controls are presented according to quartiles of plasma C-peptide concentration among controls in Table 1. In general, women with higher C-peptide levels were older than women with lower levels. Women with higher C-peptide levels also had higher BMI than those with lower levels (mean BMI for highest quartile vs. lowest: 29.0 kg/m² vs. 23.0 kg/m²). Further, women with higher C-peptide level tended to consume less alcohol (gram/day) and were less physically active (mean MET-h/week for highest quartile vs. lowest: 14.7 vs. 20.2) as compared to women with lower C-peptide level. As expected, those who provided fasting blood samples tended to have lower plasma C-peptide levels compared to those who were non-fasting. There were no other substantial differences in breast cancer risk factors across quartiles of C-peptide levels. The majority of women included in the study were white.

C-peptide and mammographic density parameters

Table 2 shows the association between quartiles of plasma C-peptide concentration, and mammographic density parameters (i.e., V measure, and square root-transformed mammographic density parameters including percent mammographic density, dense area, and non-dense area) with the lowest quartile considered as the reference category. We observed a statistically significant inverse association between plasma C-peptide concentration and percent mammographic density in the adjusted model (Q4 vs. Q1, $\beta = -1.16$, 95% CI = -1.40 to -0.91 ; *p*-trend < 0.0001). However, additional adjustment for BMI at blood collection substantially attenuated the associations (Q4 vs. Q1, $\beta = -0.12$, 95% CI = -0.36 to 0.12 ; *p*-trend = 0.009). We observed a strong positive association between C-peptide concentration and non-dense area (Q4 vs. Q1, $\beta = 3.18$, 95% CI = 2.70 – 3.54 , *p*-trend < 0.0001). After adjustment for BMI, the associations were

Table 1 | Characteristics of breast cancer controls at the time of blood collection, according to quartiles of plasma C-peptide concentration

Characteristics	bQuartile of plasma C-peptide concentration			
	Q1 N = 586	Q2 N = 571	Q3 N = 561	Q4 N = 503
Age, years, mean, (SD)	52.7 (8.2)	52.9 (8.5)	54.2 (8.1)	53.5 (8.4)
Menopausal status, N (%)				
Premenopausal	283 (48.3)	277 (48.5)	221 (39.4)	212 (42.1)
Postmenopausal	303 (51.7)	294 (51.5)	340 (60.6)	291 (57.9)
MHT status, N (%)				
Postmenopausal never used MHT	96 (16.4)	77 (13.5)	128 (22.8)	107 (21.3)
Postmenopausal formerly used MHT	69 (11.8)	69 (12.1)	76 (13.5)	65 (12.9)
Postmenopausal currently using MHT	93 (15.9)	118 (20.7)	101 (18.0)	83 (16.5)
Age at menopause ^a , years, mean (SD)	48.5 (4.7)	49.0 (4.7)	48.9 (5.2)	48.7 (4.6)
BMI, kg/m ² , mean, (SD)	23.0 (2.9)	24.3 (3.9)	26.7 (4.9)	29.0 (6.5)
Waist-hip-ratio, mean, (SD)	0.8 (0.1)	0.8 (0.1)	0.8 (0.1)	0.8 (0.1)
Age at menarche, years, mean, (SD)	12.7 (1.4)	12.6 (1.4)	12.5 (1.4)	12.3 (1.4)
Age at first birth/parity, N (%)				
Nulliparous	66 (11.3)	50 (8.8)	53 (9.4)	37 (7.4)
1 or 2 children and <25 years old	71 (12.1)	69 (12.1)	66 (11.8)	72 (14.3)
1 or 2 children and 25+ years old	163 (27.8)	130 (22.8)	138 (24.6)	126 (25.0)
3+ children and <25 years old	158 (27.0)	187 (32.7)	185 (33.0)	144 (28.6)
3+ children and 25+ years old	124 (21.2)	130 (22.8)	118 (21.0)	122 (24.3)
Alcohol consumption, gram/day, mean (SD)	8.9 (9.5)	9.3 (10.7)	8.0 (10.1)	7.7 (11.7)
Physical activity, mean (SD), MET-h/week	20.2 (24.5)	16.4 (17.5)	16.7 (19.8)	14.7 (30.2)
Benign breast diseases, N (%)				
No	468 (79.9)	475 (83.2)	457 (81.5)	413 (82.1)
Yes	118 (20.1)	96 (16.8)	104 (18.5)	90 (17.9)
Family history of breast cancer, N (%)				
No	526 (89.8)	514 (90.0)	496 (88.4)	452 (89.9)
Yes	60 (10.2)	57 (10.0)	65 (11.6)	51 (10.1)
Diabetes, N (%)				
No	568 (96.9)	564 (98.8)	548 (97.7)	493 (98.0)
Yes	18 (3.1)	7 (1.2)	13 (2.3)	10 (2.0)
Race/ethnicity, N (%)				
Non-Hispanic White	557 (95.1)	554 (95.3)	536 (95.5)	481 (95.6)
Non-Hispanic Black	5 (0.9)	5 (0.9)	2 (0.4)	2 (0.4)
Hispanic	4 (0.7)	3 (0.5)	1 (0.2)	2 (0.4)
Asian	2 (0.3)	3 (0.5)	4 (0.7)	1 (0.2)
Unknown	18 (3.1)	16 (2.8)	18 (3.2)	17 (3.4)
Fasting status, N (%)				
Yes	469 (80.0)	446 (78.1)	402 (71.7)	225 (44.7)
No	117 (20.0)	125 (21.9)	159 (28.3)	278 (55.3)

Nurses' Health Study and Nurses' Health Study II combined (n = 2221).

^aAmong menopausal women.

^bUsing batch-specific quartiles. The quartile cut points are among the controls only.

attenuated but remained statistically significant (Q4 vs. Q1, $\beta = 0.67$, 95% CI = 0.30–1.03, p -trend < 0.0001). We observed no associations between C-peptide and dense area after additional adjustment for BMI. Similar findings were observed for the V measure. The effect estimates for the association between quartiles of C-peptide concentration and mammographic percent density were further attenuated and become not statistically significant after additional adjustment for waist-to-hip-ratio and BMI quadratic term (Supplementary Table 2). Similarly, the associations between C-peptide concentration and non-dense area were also attenuated after additional adjustment for waist-to-hip-ratio but remained statistically significant (Supplementary Table 2). Stratifying analyses by menopausal status, the results were generally consistent with the results from the overall analyses (Supplementary Table 3). Additional analyses stratified by BMI, showed similar results for women with healthy BMIs (BMI < 25 kg/m²) and for overweight and obese women (BMI ≥ 25 kg/m²) (Supplementary Table 4). The sensitivity analyses among the subset of women without diabetes diagnosis (N = 2173), showed results that were not significantly different from the overall analyses (Supplementary Table 5).

C-peptide and breast cancer

Table 3 shows the association between quartiles of plasma C-peptide concentration with breast cancer. The mean years between blood drawn and breast cancer occurrence was 6.0 years (SD = 4.0). We observed a 46% increased risk of invasive breast cancer among women with the highest quartile of plasma C-peptide (adjusted OR = 1.46, 95% CI: 1.12–1.91, p -trend = 0.0009) compared to the lowest category of C-peptide. The association was slightly more pronounced for premenopausal women (adjusted OR = 1.58, 95% CI = 1.08–2.29, p -trend = 0.02), but this was not significantly different than for postmenopausal women (p -heterogeneity = 0.45). There was also no significant heterogeneity by ER status (p -heterogeneity = 0.73). The results from the sensitivity analyses after excluding those with diabetes diagnosis (controls N = 48, cases N = 49) remained largely unchanged (Supplementary Table 6). Table 4 shows the joint effect of plasma C-peptide, the measure of mammographic density parameters, and V measure on invasive breast cancer risk. We did not find evidence that mammographic density parameters and V measure modified the association between plasma C-peptide concentration and the risk of invasive breast cancer on the multiplicative scale (p for interaction ≥ 0.07 for all mammographic parameters) (Supplementary Fig. 1). Similar results were observed for pre- and postmenopausal women when stratified by menopausal status (Supplementary Tables 7 and 8).

Discussion

In this large, well-characterized, nested case-control study, we found a significant inverse association between plasma C-peptide levels and percent mammographic density which was attenuated after adjusting for BMI, WHR, and BMI quadratic (Supplementary Table 2). A positive association between C-peptide and non-dense area was seen in both pre- and postmenopausal women. We observed a positive association between plasma C-peptide and invasive breast cancer. The association between C-peptide and breast cancer did not significantly vary by mammographic density parameters and texture V measure.

There are only a few studies that have evaluated the association between plasma C-peptide level and mammographic density. Our main findings are consistent with results from previous studies^{26,27}, including our own earlier report⁴. In the Canadian cross-sectional study of 1499 healthy women (including 787 premenopausal and 804 postmenopausal women) showed a correlation between high plasma C-peptide and low breast density²⁶. However, the strength of the negative correlation was substantially reduced and was no longer significant after further adjustments for BMI and waist-to-hip ratio, suggesting that C-peptide is not associated with breast density after adjustment for adiposity²⁶. Although we adjusted for BMI and waist-to-hip-ratio in our analyses, it is possible that adiposity may still have an important impact on some of the outcomes in this study. In our study, both BMI and waist to hip ratio were self-reported. Also, the timing of the

Table 2 | Association of quartiles of C-peptide concentration, square-root transformed mammographic density parameters, and V measure among controls, Nurses' Health Study and Nurses' Health Study II combined (N = 2221)

All women	Quartile of plasma C-peptide concentration				p value of trend [§]
	Q1 N = 586	Q2 N = 571	Q3 N = 561	Q4 N = 503	
Square root mammographic percent density					
Mean (SD)	5.8 (1.9)	5.4 (1.9)	4.8 (2.0)	4.4 (2.0)	–
Multivariable adjusted ^a	Ref.	–0.34 (–0.56 to –0.13)	–0.85 (–1.08 to –0.63)	–1.16 (–1.40 to –0.91)	<0.0001
Multivariable adjusted + BMI ^b	Ref.	–0.10 (–0.30 to 0.11)	–0.25 (–0.46 to –0.03)	–0.12 (–0.36 to 0.12)	0.009
Square root dense area					
Mean (SD)	6.2 (2.1)	6.2 (2.1)	5.9 (2.5)	5.7 (2.3)	–
Multivariable adjusted ^a	Ref.	0.06 (–0.22 to 0.34)	–0.21 (–0.50 to 0.07)	–0.30 (–0.61 to –0.00)	0.005
Multivariable adjusted + BMI ^b	Ref.	0.15 (–0.12 to 0.43)	0.04 (–0.29 to 0.30)	0.06 (–0.27 to 0.39)	0.51
Square root non dense area (cm ²)					
Mean (SD)	8.7 (2.8)	9.8 (3.1)	11.0 (3.5)	12.1 (3.9)	–
Multivariable adjusted ^a	Ref.	1.01 (0.63 to 1.40)	2.06 (1.67 to 2.45)	3.18 (2.70 to 3.54)	<0.0001
Multivariable adjusted + BMI ^b	Ref.	0.44 (0.13 to 0.75)	0.61 (0.29 to 0.94)	0.67 (0.30 to 1.03)	<0.0001
All women ^c					
V measure					
Mean (SD)	0.1 (0.9)	0.0 (1.0)	–0.1 (1.0)	–0.3 (1.0)	–
Multivariable adjusted ^a	Ref.	–0.06 (–0.18 to 0.06)	–0.16 (–0.29 to –0.04)	–0.28 (–0.42 to –0.15)	<0.0001
Multivariable adjusted + BMI ^b	Ref.	0.01 (–0.10 to 0.13)	0.04 (–0.08 to 0.16)	0.05 (–0.09 to 0.19)	0.84

The quartile cut points are among the controls only.

BMI body mass index, MHT menopausal hormone therapy, SD standard deviation.

[§]p value; tests for trend were conducted using the Wald test, where the medians of the quartiles were modeled continuously.

^aAdjusted for age at blood draw (continuous), alcohol consumption gram/day (continuous), the batch of mammography density reading (categorical), age at menarche (continuous), age at first birth/parity (nulliparous; 1 or 2 children and <25 years old; 1 or 2 children and 25+ years old; 3+ children and <25 years old; or 3+ children and 25+ years old), menopausal status (premenopausal; postmenopausal and MHT never users; postmenopausal and former MHT users; postmenopausal and current MHT users), physical activity (continuous), benign breast diseases (yes, no), family history of breast cancer (yes, no), fasting status, diabetes status (yes, no) at blood drawn.

^bAdjusted additionally for BMI (continuous).

^cAmong women with V measure (N = 2011).

BMI was assessment was from the questionnaire closest to the mammogram, and we only had one assessment of waist to hip ratio, therefore there are multiple sources of measurement error in these measurements. Given the measurement error, we expect that there is residual confounding by adiposity. If we were better able to adjust for adiposity at the time of the mammogram, these associations may have been further attenuated.

Consistent with other observational studies^{4,5,8} and a meta-analysis¹⁵, we found a positive association between plasma C-peptide level and the risk of invasive breast cancer among both pre- and postmenopausal women. A meta-analysis of 9 observational studies (5 case-control and 4 nested-case control studies within prospective studies) reported an overall increased risk of breast cancer by 26% (pooled risk ratio = 1.26, 95% CI 1.06–1.48¹⁵, comparing those in the upper categories of C-peptide/insulin to those in the lowest. However, there was significant heterogeneity by study design—the pooled estimate of the nested-case control studies was a non-significant risk ratios of 1.04 while the case-control studies gave a summary estimate of 1.81 (95% CI 1.36–2.40). The risk ratios were adjusted for BMI in all studies except one. A case-control study of 143 pairs matched by age and menopausal status from the Shanghai Breast Cancer Study cohort reported a dose-response relationship between C-peptide and risk of breast cancer with nearly a 3-fold increased risk among women with the highest C-peptide levels independent of BMI and menopausal status⁸.

C-peptide is a biological by-product of insulin processing and is considered to be a marker of hyperinsulinemia. Due to its stability in plasma, C-peptide is clinically used to evaluate pancreatic beta-cell function. Two hypotheses have been proposed to explore the underlying biological mechanisms of the association between elevated C-peptide/insulin levels and increased risk of breast cancer. First, C-peptide may directly potentiate

the insulin receptor. Second, C-peptide may increase the concentration of bioavailability of sex hormones, and thereby influence the action of insulin on breast cancer cell growth directly and/or indirectly³². Laboratory studies showed that insulin, like its associated growth factors, may have mitotic effects on both normal and neoplastic breast epithelial cells^{44,45}.

Our study has several strengths. The availability of high-quality, prospectively collected information on exposure, covariate, and outcome data. Detailed information on lifestyle, demographic, and reproductive factors within both cohorts enabled us to account for potential confounding effects of lifestyle factors (e.g., physical activity and alcohol consumption), reproductive, and adiposity factors on the measured associations. Also, we were able to evaluate associations by tumor ER subtype, and menopausal status. Further, to our knowledge, this study is the first to report, whether mammographic density parameters and V measure modify the association between plasma C-peptide level and risk of breast cancer. There were also limitations in our study. Although we had comprehensive information on breast cancer risk factors in this study which allowed for adjustment for potential confounders, the potential for residual confounding especially with respect to adiposity is a concern. Our findings have to be interpreted with caution because they are based on a single measurement of plasma C-peptide, which might not reflect a full picture of insulin resistance and elevation of insulin resistance markers. Additionally, one-time measurement of C-peptide might reduce our ability to evaluate associations between long-term circulating concentrations of C-peptide and mammographic features. However, C-peptide levels have previously been reported to be relatively stable with a moderately high within-person coefficient of variation over 4 years ($r = 0.57$)⁴⁶. Future research looking at various parameters of insulin resistance including glucose, homeostatic model assessment

Table 3 | Odds ratios (95% confidence intervals) of quartiles C-peptide concentration and breast cancer risk in Nurses' Health Study and Nurses' Health Study II combined

	Quartile of plasma C-peptide concentration				p value [§]	p-heterogeneity
	Q1	Q2	Q3	Q4		
All women						
Invasive breast cancer (1023/2221) ^a						
Multivariate adjusted*	1.00 (Ref.)	1.10 (0.90–1.41)	0.96 (0.76–1.22)	1.40 (1.09–1.78)	0.002	
Multivariate adjusted + BMI	1.00 (Ref.)	1.12 (0.88–1.41)	1.00 (0.77–1.27)	1.46 (1.12–1.91)	0.0009	
Invasive and in situ breast cancers (1260/2221)						
Multivariate adjusted*	1.00 (Ref.)	1.09 (0.89–1.35)	1.08 (0.86–1.33)	1.32 (1.05–1.65)	0.09	
Multivariate adjusted + BMI	1.00 (Ref.)	1.11 (0.90–1.37)	1.12 (0.90–1.41)	1.46 (1.14–1.87)	0.05	
ER (+) invasive breast cancer (913/2221)						
Multivariate adjusted*	1.00 (Ref.)	1.12 (0.88–1.42)	1.06 (0.82–1.35)	1.45 (1.12–1.86)	0.005	0.73 ^a
Multivariate adjusted + BMI	1.00 (Ref.)	1.14 (0.90–1.46)	1.11 (0.86–1.44)	1.57 (1.20–2.07)	0.002	
ER (-) invasive breast cancer (225/2221)						
Multivariate adjusted*	1.00 (Ref.)	1.12 (0.72–1.74)	1.13 (0.72–1.77)	1.60 (1.03–2.52)	0.07	
Multivariate adjusted + BMI	1.00 (Ref.)	1.15 (0.74–1.78)	1.20 (0.74–1.89)	1.77 (1.09–2.90)	0.05	
Premenopausal women						
Invasive breast cancer (341/993)						
Multivariate adjusted*	1.00 (Ref.)	1.21 (0.84–1.73)	1.12 (0.76–1.66)	1.49 (1.01–2.22)	0.02	0.45 ^b
Multivariate adjusted + BMI	1.00 (Ref.)	1.12 (0.81–1.55)	1.41 (1.00–1.99)	1.58 (1.08–2.29)	0.02	
ER (+) invasive breast cancer (334/993)						
Multivariate adjusted*	1.00 (Ref.)	1.19 (0.83–1.72)	1.22 (0.83–1.79)	1.53 (1.03–2.27)	0.008	
Multivariate adjusted + BMI	1.00 (Ref.)	1.23 (0.85–1.77)	1.33 (0.89–1.98)	1.76 (1.16–2.70)	0.002	
Postmenopausal women [†]						
Invasive breast cancer (682/1228)						
Multivariate adjusted*	1.00 (Ref.)	1.01 (0.74–1.37)	0.88 (0.65–1.20)	1.37 (1.00–1.88)	0.004	0.34 ^c
Multivariate adjusted + BMI	1.00 (Ref.)	1.01 (0.74–1.37)	0.87 (0.63–1.20)	1.36 (0.96–1.92)	0.004	
ER (+) invasive breast cancer (579/1228)						
Multivariate adjusted*	1.00 (Ref.)	1.03 (0.74–1.43)	0.96 (0.70–1.33)	1.40 (1.00–1.95)	0.04	
Multivariate adjusted + BMI	1.00 (Ref.)	1.03 (0.74–1.44)	0.97 (0.69–1.36)	1.41 (0.98–2.03)	0.05	

*Adjusted for age, age at menarche (continuous), age at first birth/parity (nulliparous; 1 or 2 children and <25 years old; 1 or 2 children and 25+ years old; 3+ children and <25 years old; or 3+ children and 25+ years old), physical activity (continuous), alcohol consumption gram/day (continuous), menopausal status (premenopausal; postmenopausal and MHT never users; postmenopausal and former MHT users; postmenopausal and current MHT users), benign breast diseases (yes, no), family history of breast cancer (yes, no), fasting status, diabetes status (yes, no) blood collection, and matching factors.

[†]Adjusted for menopausal hormone therapy use (never/former/current).

^aNumber of women with invasive ductal carcinoma (N = 868), invasive lobular carcinoma (N = 124), and with both (N = 31).

[§]p value; tests for trend were conducted using the Wald test, where the medians of the quartiles were modeled continuously.

^bp value of heterogeneity between ER-positive and ER-negative.

^cp value of heterogeneity between premenopausal invasive breast cancer and postmenopausal invasive breast cancer.

^dp value of heterogeneity between premenopausal ER-positive and postmenopausal ER-positive breast cancer.

(HOMA), and hemoglobin A1C (HbA1c) and breast cancer risk is clearly warranted. Additionally, future cohort studies are needed to collect real-time data such as glucose and insulin sensors and examine those data in relation to breast cancer risk. Nevertheless, it is important to mention that previous study showed insulin levels have been fairly consistent over time (the intraclass correlation coefficient = 0.70 for samples collected a year apart⁷). This is especially relevant when considering the cross-sectional associations between C-peptide and mammographic parameters; the lack of positive association in this study does not preclude C-peptide levels earlier in life having a different association with mammographic features. In this study, C-peptide was measured over 10 batches; therefore, some degree of measurement error and information bias is inevitable. Nevertheless, the percent CV for C-peptide measurement ranged from 5.3% to 9.0% (mean = 6.7%) over all assay batches. Additionally, the information bias is more likely nondifferential because women were not aware of their mammographic

feature measurement or the potential association of C-peptide on breast cancer risk. If anything, our estimates could therefore be diluted. Finally, the study population consists largely of white nurses, potentially limiting generalizability, and additional studies in more diverse populations are needed to further confirm our results. In conclusion, the results of this large prospective analysis suggest a positive association between C-peptide concentration, a proxy measurement of insulin secretion, and the risk of breast cancer in pre- and postmenopausal women. Further, mammographic density parameters and V texture feature measures do not seem to modify the association between C-peptide and the risk of breast cancer. These findings suggest that C-peptide/insulin level influences breast cancer risk through other mechanisms than mammographic density and mammographic texture feature. High levels of C-peptide/insulin have been shown to be a risk factor for breast cancer, and we found no evidence that mammographic density parameters modify this association.

Table 4 | Odds ratios (95% confidence intervals) of cross-classified quartiles of C-peptide concentration and quartiles of mammographic density parameters on invasive breast cancer risk

	^a Quartile of plasma C-peptide concentration				<i>p</i> value ^b
	Q1	Q2	Q3	Q4	
Percent MD (%)					0.61
Q1, <14.78	1.00 (Ref.)	1.07 (0.60–2.00)	0.94 (0.54–1.67)	1.27 (0.73–2.25)	
Q2, ≥14.78 to <29.12	1.09 (0.61–2.00)	1.14 (0.53–2.44)	1.44 (0.68–3.02)	1.50 (0.73–3.07)	
Q3, ≥29.13 to <47.69	2.21 (1.03–3.15)	0.80 (0.39–1.65)	0.72 (0.34–1.47)	1.00 (0.47–2.00)	
Q4, ≥47.69	1.78 (1.03–3.15)	1.34 (0.65–2.75)	1.42 (0.68–2.93)	1.41 (0.66–2.97)	
Dense area (cm ²)					0.96
Q1, <22.42	1.00 (Ref.)	1.24 (0.72–2.12)	1.05 (0.63–1.77)	1.62 (0.96–2.78)	
Q2, ≥22.42 to <37.56	1.65 (1.00–2.74)	0.80 (0.39–1.58)	0.84 (0.42–1.70)	0.64 (0.32–1.28)	
Q3, ≥37.56 to <58.57	1.70 (1.02–2.84)	0.93 (0.47–1.89)	1.09 (0.53–2.19)	1.08 (0.54–2.14)	
Q4, ≥58.57	2.06 (1.24–3.47)	0.88 (0.44–1.75)	0.93 (0.46–1.85)	0.91 (0.45–1.81)	
Non-dense area (cm ²)					0.50
Q1, <56.48	1.00 (Ref.)	1.50 (1.00–2.24)	1.16 (0.70–2.24)	1.57 (0.92–2.65)	
Q2, ≥56.48 to <97.21	0.83 (0.54–1.27)	0.67 (0.37–1.21)	1.11 (0.56–2.21)	0.94 (0.46–1.91)	
Q3, ≥97.21 to <155.68	0.68 (0.43–1.07)	0.73 (0.39–1.37)	0.76 (0.38–1.53)	1.13 (0.56–2.28)	
Q4, ≥155.68	0.85 (0.47–1.56)	0.52 (0.25–1.10)	0.59 (0.27–1.30)	0.61 (0.28–1.34)	
V measure					0.07
Q1, <−0.69	1.00 (Ref.)	1.00 (0.49–1.69)	1.45 (0.85–2.54)	1.46 (0.87–2.51)	
Q2, ≥−0.69 to <0.004	1.28 (0.72–2.31)	1.17 (0.51–2.66)	0.90 (0.42–1.87)	1.16 (0.56–2.39)	
Q3, ≥0.004 to ≤0.66	1.38 (0.80–2.41)	1.09 (0.50–2.31)	0.71 (0.34–1.50)	0.85 (0.41–1.74)	
Q4, >0.66	1.37 (0.79–2.42)	1.40 (0.64–3.07)	1.04 (0.50–2.15)	1.19 (0.56–2.50)	

^aAdjusted age, and BMI at blood collection (continuous), at menarche (continuous), age at first birth/parity (nulliparous; 1 or 2 children and <25 years old; 1 or 2 children and 25+ years old; 3+ children and <25 years old; or 3+ children and 25+ years old), physical activity (continuous), alcohol consumption gram/day (continuous), menopausal status (premenopausal; postmenopausal and MHT never users; postmenopausal and former MHT users; postmenopausal and current MHT users), benign breast diseases (yes, no), family history of breast cancer (yes, no), fasting status, diabetes status (yes, no) at blood collection, and waist to hip circumference ratio (continuous) at blood collection and matching factors.

^bMultiplicative interaction was tested with a likelihood ratio test through comparing 2 generalized linear models with and without the interaction term.

Data availability

The data that support the findings of this study are available from Nurses' Health Study, but restrictions apply to the availability of these data, which were used under license for the current study, and so are not publicly available. Data are however available from the authors upon reasonable request and with permission.

Code availability

The underlying code for this study is not publicly available but may be made available to qualified researchers on reasonable request from the corresponding author.

Received: 25 September 2023; Accepted: 7 October 2024;

Published online: 17 October 2024

References

- Venugopal, S. K., Mowery, M. L. & Jialal, I. *Biochemistry, C Peptide. StatPearls* (StatPearls Publishing LLC, 2024).
- Leighton, E., Sainsbury, C. A. & Jones, G. C. A practical review of C-peptide testing in diabetes. *Diabetes Ther.* **8**, 475–487 (2017).
- Li, M. et al. Association between serum insulin and C-peptide levels and breast cancer: an updated systematic review and meta-analysis. *Front. Oncol.* **10**, 553332 (2020).
- Ahern, T. P. et al. Plasma C-peptide, mammographic breast density, and risk of invasive breast cancer. *Cancer Epidemiol. Biomark. Prev.* **22**, 1786–1796 (2013).
- Bruning, P. F. et al. Insulin resistance and breast-cancer risk. *Int. J. Cancer* **52**, 511–516 (1992).
- Del Giudice, M. E. et al. Insulin and related factors in premenopausal breast cancer risk. *Breast Cancer Res. Treat.* **47**, 111–120 (1998).
- Muti, P. et al. Fasting glucose is a risk factor for breast cancer: a prospective study. *Cancer Epidemiol. Biomark. Prev.* **11**, 1361–1368 (2002).
- Yang, G. et al. Population-based, case-control study of blood C-peptide level and breast cancer risk. *Cancer Epidemiol. Biomark. Prev.* **10**, 1207–1211 (2001).
- Eliassen, A. H., Tworoger, S. S., Mantzoros, C. S., Pollak, M. N. & Hankinson, S. E. Circulating insulin and c-peptide levels and risk of breast cancer among predominately premenopausal women. *Cancer Epidemiol. Biomark. Prev.* **16**, 161–164 (2007).
- Mink, P. J., Shahar, E., Rosamond, W. D., Alberg, A. J. & Folsom, A. R. Serum insulin and glucose levels and breast cancer incidence: the atherosclerosis risk in communities study. *Am. J. Epidemiol.* **156**, 349–352 (2002).
- Schairer, C. et al. Serum concentrations of IGF-I, IGFBP-3 and c-peptide and risk of hyperplasia and cancer of the breast in postmenopausal women. *Int. J. Cancer* **108**, 773–779 (2004).
- Toniolo, P. et al. Serum insulin-like growth factor-I and breast cancer. *Int. J. Cancer* **88**, 828–832 (2000).
- Yu, H. & Rohan, T. Role of the insulin-like growth factor family in cancer development and progression. *J. Natl Cancer Inst.* **92**, 1472–1489 (2000).
- Gunter, M. J. et al. Insulin, insulin-like growth factor-I, and risk of breast cancer in postmenopausal women. *J. Natl Cancer Inst.* **101**, 48–60 (2009).
- Pisani, P. Hyper-insulinaemia and cancer, meta-analyses of epidemiological studies. *Arch. Physiol. Biochem.* **114**, 63–70 (2008).

16. Belfiore, A. & Frasca, F. IGF and insulin receptor signaling in breast cancer. *J. Mammary Gland Biol. Neoplasia* **13**, 381–406 (2008).
17. Boyd, N. F. et al. Breast tissue composition and susceptibility to breast cancer. *J. Natl Cancer Inst.* **102**, 1224–1237 (2010).
18. Vachon, C. M. et al. Mammographic density, breast cancer risk and risk prediction. *Breast Cancer Res.* **9**, 217 (2007).
19. Byng, J. W. et al. Symmetry of projection in the quantitative analysis of mammographic images. *Eur. J. Cancer Prev.* **5**, 319–327 (1996).
20. McCormack, V. A. & dos Santos Silva, I. Breast density and parenchymal patterns as markers of breast cancer risk: a meta-analysis. *Cancer Epidemiol. Biomark. Prev.* **15**, 1159–1169 (2006).
21. Malkov, S. et al. Mammographic texture and risk of breast cancer by tumor type and estrogen receptor status. *Breast Cancer Res.* **18**, 122 (2016).
22. Manduca, A. et al. Texture features from mammographic images and risk of breast cancer. *Cancer Epidemiol. Biomark. Prev.* **18**, 837–845 (2009).
23. Nielsen, M. et al. Mammographic texture resemblance generalizes as an independent risk factor for breast cancer. *Breast Cancer Res.* **16**, R37 (2014).
24. Wanders, J. O. P. et al. The combined effect of mammographic texture and density on breast cancer risk: a cohort study. *Breast Cancer Res.* **20**, 36 (2018).
25. Heine, J. J. et al. A novel automated mammographic density measure and breast cancer risk. *J. Natl Cancer Inst.* **104**, 1028–1037 (2012).
26. Diorio, C. et al. Levels of C-peptide and mammographic breast density. *Cancer Epidemiol. Biomark. Prev.* **14**, 2661–2664 (2005).
27. Wolin, K. Y. et al. Associations of physical activity, sedentary time, and insulin with percent breast density in Hispanic women. *J. Womens Health* **16**, 1004–1011 (2007).
28. Nurses' Health Study. History of Nurses' Health Study. Developed by Health Communication Core. (2016). Access (2024). [<https://nurseshealthstudy.org/about-nhs/history>].
29. Hankinson, S. E. et al. Alcohol, height, and adiposity in relation to estrogen and prolactin levels in postmenopausal women. *J. Natl Cancer Inst.* **87**, 1297–1302 (1995).
30. Eliassen, A. H. et al. Endogenous steroid hormone concentrations and risk of breast cancer among premenopausal women. *J. Natl Cancer Inst.* **98**, 1406–1415 (2006).
31. Tamimi, R. M., Byrne, C., Colditz, G. A. & Hankinson, S. E. Endogenous hormone levels, mammographic density, and subsequent risk of breast cancer in postmenopausal women. *J. Natl Cancer Inst.* **99**, 1178–1187 (2007).
32. Pollak, M. Insulin and insulin-like growth factor signalling in neoplasia. *Nat. Rev. Cancer* **8**, 915–928 (2008).
33. Bertrand, K. A. et al. Urinary estrogens and estrogen metabolites and mammographic density in premenopausal women. *Breast Cancer Res. Treat.* **136**, 277–287 (2012).
34. Colditz, G. A. & Hankinson, S. E. The Nurses' Health Study: lifestyle and health among women. *Nat. Rev. Cancer* **5**, 388–396 (2005).
35. Rice, M. S., Rosner, B. A. & Tamimi, R. M. Percent mammographic density prediction: development of a model in the nurses' health studies. *Cancer Causes Control* **28**, 677–684 (2017).
36. Rice, M. S. et al. Mammographic density and breast cancer risk: a mediation analysis. *Breast Cancer Res.* **18**, 94 (2016).
37. Rosner, B., Cook, N., Portman, R., Daniels, S. & Falkner, B. Determination of blood pressure percentiles in normal-weight children: some methodological issues. *Am. J. Epidemiol.* **167**, 653–666 (2008).
38. Harris, J. R., Lippman, M. E., Veronesi, U. & Willett, W. Breast cancer (1). *N. Engl. J. Med.* **327**, 319–328 (1992).
39. Stampfer, M. J. et al. Postmenopausal estrogen therapy and cardiovascular disease. Ten-year follow-up from the nurses' health study. *N. Engl. J. Med.* **325**, 756–762 (1991).
40. Tamimi, R. M., Hankinson, S. E., Chen, W. Y., Rosner, B. & Colditz, G. A. Combined estrogen and testosterone use and risk of breast cancer in postmenopausal women. *Arch. Intern. Med.* **166**, 1483–1489 (2006).
41. Baglietto, L. et al. Associations of mammographic dense and nondense areas and body mass index with risk of breast cancer. *Am. J. Epidemiol.* **179**, 475–483 (2014).
42. Soguel, L., Durocher, F., Tchernof, A. & Diorio, C. Adiposity, breast density, and breast cancer risk: epidemiological and biological considerations. *Eur. J. Cancer Prev.* **26**, 511–520 (2017).
43. Cohen, J. F. et al. Cochran's Q test was useful to assess heterogeneity in likelihood ratios in studies of diagnostic accuracy. *J. Clin. Epidemiol.* **68**, 299–306 (2015).
44. Kaaks, R. Nutrition, insulin, IGF-1 metabolism and cancer risk: a summary of epidemiological evidence. *Novartis Found. Symp.* **262**, 247–260 (2004).
45. Rose, D. P., Komninou, D. & Stephenson, G. D. Obesity, adipocytokines, and insulin resistance in breast cancer. *Obes. Rev.* **5**, 153–165 (2004).
46. Ma, J. et al. A prospective study of plasma C-peptide and colorectal cancer risk in men. *J. Natl Cancer Inst.* **96**, 546–553 (2004).

Acknowledgements

The authors thank the participants and staff of the Nurses' Health Study and Nurses' Health Study II for their valuable contributions. The authors would like to acknowledge the contribution to this study from central cancer registries supported through the Centers for Disease Control and Prevention's National Program of Cancer Registries (NPCR) and/or the National Cancer Institute's Surveillance, Epidemiology, and End Results (SEER) Program. Central registries may also be supported by state agencies, universities, and cancer centers. Participating central cancer registries include the following: Alabama, Alaska, Arizona, Arkansas, California, Colorado, Connecticut, Delaware, Florida, Georgia, Hawaii, Idaho, Indiana, Iowa, Kentucky, Louisiana, Massachusetts, Maine, Maryland, Michigan, Mississippi, Montana, Nebraska, Nevada, New Hampshire, New Jersey, New Mexico, New York, North Carolina, North Dakota, Ohio, Oklahoma, Oregon, Pennsylvania, Puerto Rico, Rhode Island, Seattle SEER Registry, South Carolina, Tennessee, Texas, Utah, Virginia, West Virginia, Wyoming. This work was supported by the National Cancer Institute at the National Institutes of Health [CA131332, CA124865, CA175080 to R.M.T., UM1 CA186107 and U01 CA200464 to J.H.], Avon Foundation for Women, Susan G. Komen for the Cure® (to R.M.T.), and Breast Cancer Research Foundation.

Author contributions

Shadi Azam: research design, analyses, writing of the manuscript, critical input into the writing of the manuscript, and reading and approval of the final manuscript. Rulla M. Tamimi: conceptualization, project supervision, and critical input into the writing of the manuscript, and reading and approval of the final manuscript. All authors have made substantial contributions to the interpretation of data, revisions to the manuscript for important intellectual content, writing—original draft, and review and approval of the final version.

Competing interests

The authors declare no competing interests.

Additional information

Supplementary information The online version contains supplementary material available at <https://doi.org/10.1038/s41523-024-00702-x>.

Correspondence and requests for materials should be addressed to Shadi Azam.

Reprints and permissions information is available at <http://www.nature.com/reprints>

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Open Access This article is licensed under a Creative Commons Attribution-NonCommercial-NoDerivatives 4.0 International License, which permits any non-commercial use, sharing, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if you modified the licensed material. You do not have permission under this licence to share adapted material derived from this article or parts of it. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by-nc-nd/4.0/>.

© The Author(s) 2024