RESEARCH

Breast Cancer Research



Open Access

Associations of circulating insulin-like growth factor-1 and insulin-like growth factor binding protein-3 with the expression of stem cell markers in benign breast tissue

Lusine Yaghjyan^{1*}, Yujing J. Heng², Brian R. Sardella², Divya Murthy³, Matt B. Mahoney³, Bernard Rosner³, Kornelia Polyak⁴, Maisey Ratcliff¹ and Rulla M. Tamimi⁵

Abstract

Background The insulin-like growth factor (IGF) pathway is implicated in a naturally occurring process of tissue remodeling during which cells acquire stem cell-like characteristics. We examined associations of circulating IGF-1 and IGF binding protein-3 (IGFBP-3) with expression of CD44, CD24, and ALDH1A1 stem cell markers in benign breast biopsies.

Methods This study included 151 cancer-free women with incident biopsy-confirmed benign breast disease and blood samples within the Nurses' Health Study II. The data on reproductive and other BCa risk factors were obtained from biennial questionnaires. Immunohistochemistry (IHC) was done on tissue microarrays. For each core, the IHC expression was assessed using QuPath, and expressed as % of cells that stain positively for a specific marker out of the total cell count. Generalized linear regression was used to examine the associations of plasma IGF-I and IGFBP-3 (continuous log-transformed and quartiles) with log-transformed expression of each marker (in epithe-lium and stroma), adjusted for BCa risk factors.

Results In multivariate analysis, continuous circulating IGF-1 and IGFBP-3 measures were not associated with the continuous expression of any of the markers in the epithelium or stroma. Women whose IGFBP-3 levels were in the top quartile appeared to have lower expression of stromal CD24 compared to those in the lowest quartile ($\beta = -0.38$, 95% CI - 0.69, - 0.08, *p*-trend = 0.06).

Conclusions Higher circulating IGFBP-3 levels were associated with lower stromal CD24 expression in benign breast tissue. Our findings provide indirect evidence of the inducing effect of IGF pathway on epithelial-to-mesenchymal transitions and stem cell activity in the breast.

Keywords Benign breast disease, Breast cancer risk, Stem cell markers

*Correspondence: Lusine Yaghjyan Iyaghjyan@ufl.edu Full list of author information is available at the end of the article



© The Author(s) 2025. **Open Access** This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, 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 changes were made. 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/4.0/.

Background

Breast stem cells play a critical role in sustaining breast tissue architecture throughout the woman's life [1, 2]. However, potentially limitless self-renewal capacity and high susceptibility to various endogenous and exogenous mutagenic insults increase the chances of their tumorigenic transformation [1, 3]. According to the stem cell hypothesis of breast carcinogenesis, breast cancer development might be directly related to the size and mitotic activity of the stem cell pool in the breast [4]. Further, in the breast, stem cells are the only cell subpopulation that can accumulate all the oncogenic alterations [1].

Insulin-like growth factors (IGF) play an important role in the structure of breast lobules and lobular involution [5]. Epidemiologic and animal studies suggest an increase in breast cancer risk with elevated circulating IGF-1 and insulin-like growth factor-binding protein 3 (IGFBP-3) which are also associated with a number of breast cancer risk factors, including birthweight, mammographic density, benign breast disease (BBD) as well as BBD to tumor progression [5-11]. IGF-1 system plays an important role in breast carcinogenesis [9, 12] by regulating the epithelial-to-mesenchymal transition (EMT, a natural process for tissue remodeling and wound healing), and stem cell-related processes across several tissues [13, 14]. An overlap between EMT and stem cell mechanisms is supported by the evidence that cells undergoing EMT acquire stem cell-like characteristics including self-renewal, gain of specific gene expression patterns, and ability to initiate tumorigenesis [13, 15]. However, the relationship between circulating IGF pathway markers and stem cells has never been investigated. To address this knowledge gap, we examined associations of IGF-1 and IGFBP-3 with well-characterized stem cell markers CD44, CD24, and ALDH1A1. There are no universal markers for breast cancer stem cells (and normal mammary stem cells), but CD44, CD24, and ALDH1A1 stem cell markers remain the most accepted markers that have been validated and linked to younger age at diagnosis, higher odds of unfavorable tumor characteristics, and poor prognosis and chemotherapy resistant breast cancer [16-24]. Accumulating reports show that breast cancer stem cells with co-stained CD44⁺/CD24⁻ or CD44⁺/ CD24^{low} and ALDH1⁺ or ALDH1^{high} are responsible for tumor initiation, progression, metastasis, and drug resistance [16-24]. In line with this evidence, we hypothesized that higher levels of IGF-1 and IGFBP-3 would be positively associated with the expression of CD44 and ALDH1A1, and inversely associated with CD24 (consistent with associations of CD44high,

ALDH1A1high, and CD24low with unfavorable tumor features).

Materials and methods

Study population

Our analysis included women with incident biopsyconfirmed benign breast disease (BBD) in the Nurses' Health Study II (NHSII) cohort. NHSII followed registered nurses in the United States who were 25-42 years old at enrollment. After administration of the initial 1989 NHS II questionnaire, the information on breast cancer risk factors (body mass index [BMI], reproductive history, menopausal hormone therapy [MHT] use, and alcohol use) and any diagnoses of cancer or other diseases (including BBD) was updated through biennial questionnaires which were then confirmed via medical record review [25]. The initial questionnaire and all subsequent biennial questionnaires asked participants to report any diagnosis of BBD and to indicate whether it was confirmed by biopsy or aspiration. Details of this incident BBD study and the BBD assessment have been previously described [12, 26]. We obtained BBD pathology records and archived biopsy specimens for all BBD cases from their hospital pathology departments. Women were excluded if they had evidence of in situ or invasive carcinoma or unknown lesion type at the time of benign breast biopsy.

Between 1996 and 1999, 29,611 NHSII participants who were 32–45 years old provided a blood sample [27, 28]. In brief, premenopausal NHSII participants who were not using any hormones, were not pregnant, or have not breastfed in the previous 6 months, provided a 30-ml blood sample collected 7-9 days before the anticipated start of their next menstrual cycle (luteal blood draw). NHSII women who were ineligible to provide timed samples (i.e., perimenopausal, postmenopausal, had a simple hysterectomy, currently used any hormones, or declined to give timed samples) provided a single 30-ml blood sample (referred to as "untimed" samples, 22.5% of women in our final study sample). For both luteal and untimed samples, women shipped the blood to the Brigham and Women's/Harvard Cohorts Biorepository laboratory, with an ice-pack, via overnight courier, where the samples were processed, separated into plasma, red blood cell, and white blood cell components, and aliquoted into labeled cryotubes. All samples have been stored in the vapor phase of continuously monitored liquid nitrogen freezers (below -130 °C) since collection.

In the current analysis, we included 151 women who had complete data on breast cancer factors, IGF measures, and staining results for stem cell markers. All women in our sample were diabetes-free at the time of the blood collection. The study protocol was approved by the institutional review boards (IRB) of the Brigham and Women's Hospital and Harvard T.H. Chan School of Public Health, and those of participating registries as required, and University of Florida IRB. Consent was obtained or implied by return of questionnaires.

Benign breast biopsy confirmation and BBD subtypes

Hematoxylin and eosin (H&E) breast tissue slides were retrieved for biopsy-confirmed BBD patients who gave permission to review their biopsy records. The slides were previously independently reviewed by one of three pathologists in a blinded fashion, i.e. the evaluating pathologists were blinded to type of BBD noted on the original diagnosis [29, 30]. Any slide identified as having either questionable atypia or atypia was jointly reviewed by two pathologists [29, 30]. Each benign breast biopsy was classified according to the categories of Page et al. [31] as non-proliferative, proliferative without atypia, or atypical hyperplasia (ductal or lobular hyperplasia) [12].

Tissue microarray (TMA) construction of BBD samples

After centralized review of H&E stained slides, we retrieved archived FFPE benign breast biopsy blocks for participants. H&E sections of the corresponding FFPE tissue blocks were re-reviewed by a single pathologist to identify areas of benign proliferative lesions and normal terminal duct-lobular units (TDLUs), and to identify the areas from which the cores for the TMAs would be taken. Normal TDLUs were regions of histologically normal tissue that may or may not be adjacent to benign lesions (e.g., atypical ductal hyperplasia, usual ductal hyperplasia) [12]. TMAs were constructed at the Dana Farber/ Harvard Cancer Center (DF/HCC) Tissue Microarray Core Facility by obtaining 0.6-mm cores from benign lesions and TDLUs. For each woman, up to 3 cores of normal TDLU were included in the TMA blocks. We previously evaluated our TMA construction methods and confirmed a high success rate (76%) of capturing normal TDLUs in these TMA blocks [32].

Immunohistochemistry (IHC) for stem cell markers

The expression of the stem cell markers was evaluated by semi-automated IHC technique that allows the quantification of markers' expression levels and localization of the target signal to specific cells/structures. For each of the three markers one 5- μ m paraffin section was cut from a single TMA block and then stained at the University of Florida Pathology Core Lab on DAKO AutostainerPlus according to the previously standardized protocol with commercial antibodies (CD44 [DAKO] 1:25 dilution; CD24 [Invitrogen] 1:200 dilution and ALDH1A1 [Abcam] 1:300 dilution) [33–35]. Briefly, slides were de-paraffinized with xylene Page 3 of 8

and re-hydrated through decreasing concentrations of ethanol to water, including an intermediate step to quench endogenous peroxidase activity (3% hydrogen peroxide in methanol) and transferred to 1X TBS-T (Tris-buffered saline-Tween). For heat-induced antigen retrieval, sections were heated in a steamer while submerged in Citra (Biogenex, Fremont, CA) or Trilogy (Cell Marque, Rocklin, CA) for 30 min. Next, slides were (1) rinsed in 1XTBS-T and incubated with a universal protein blocker Sniper (Biocare Medical, Walnut Creek, CA) for 10 (for CD44 and ALDH1A1) or 15 min (for CD24); (2) rinsed in 1XTBS-T and co-incubated in primary antibody ALDH1A1, CD24, or CD44 for 1 h; and (3) rinsed in 1XTBS-T followed by application of conjugated secondary antibody (Mach 2 goat antirabbit horse [or mouse] radish peroxidase-conjugated, Biocare Medical, Walnut Creek, CA) for 30 min. Detection of antibodies was achieved by incubating slides in 3'3' diaminobenzidine (Vector Laboratories Inc., Burlingame, CA) for 4 min. Slides were counterstained with hematoxylin (Biocare Medical, Walnut Creek, CA) 1:10 for 3 min and mounted with Cytoseal XYL (Richard-Allen Scientific, Kalamazoo, MI). The laboratory implemented standard quality control procedures.

Image analysis

Stained TMA sections were digitized at 40×using the PhenoImager HT (Akoya Biosciences, Marlborough, MA). QuPath v0.5.0 was used to quantify the immunoreactivity of the markers [36]. For each marker, we randomly selected 12 tissue cores of variable staining intensities on a TMA to train tissue segmentation into epithelium and stroma, and determine the minimum intensity to score a cell as immunoreactive [37]. We used random forest object classifier and the "positive cell detection" command to optimize cell parameters, intensity thresholds for hematoxylin and cytoplasm (mean optical density), and used the default values for all other parameters.

We focused the current analysis on the expression of stem cell markers in normal TDLU cores for the following reasons: (1) we specifically targeted normal TDLUs in construction of these TMAs and thus the number of women with benign lesion cores was smaller and would not allow to draw meaningful conclusions; (2) in our earlier reliability study, we observed higher heterogeneity within benign lesion cores as they were represented by various lesion types [33]; and (3) we were interested in the underlying changes in the breast tissue happening early in the process of breast carcinogenesis and thus normal TDLUs were more relevant to address our research questions.

IGF-1 and IGFBP-3 laboratory assays

IGF-1 and IGFBP-3 levels were assayed in two batches for luteal and untimed samples in the NHSII by ELISA after acid extraction at the Department of Medicine and Oncology at McGill University, using reagents from Diagnostic Systems Laboratory (Webster, TX, USA). The mean coefficients of variation were 3.5 and 2.8% for IGF-1 and 1.6 and 3.7% for IGFBP-3 [28].

Covariate information

Information on breast cancer risk factors was obtained from the biennial questionnaires closest to the biopsy date. Women were considered to be postmenopausal if they reported: (1) no menstrual periods within the 12 months before biopsy with natural menopause, (2) bilateral oophorectomy, or (3) hysterectomy with one or both ovaries retained, and were 54 years or older for ever smokers or 56 years or older for never smokers [38, 39].

Statistical analysis

To account for batch-to-batch variation, IGF-1 and IGFBP-3 levels were recalibrated to have a comparable distribution to an average batch according to methods outlined by Rosner et al. [40]. We used generalized linear regression to examine the associations of plasma concentrations of IGF-I and IGFBP-3 (continuous logtransformed and batch-specific quartiles) with log-transformed expression of each marker (in epithelium and stroma, measured as weighted [by the number of cells] average across available cores for a woman), adjusted for the following covariates: (1) Model 1: age (continuous), BMI (continuous); (2) Model 2 (full): age (continuous), BMI (continuous), a family history of breast cancer (yes/no), menopausal status (premenopausal, postmenopausal, unknown menopausal status), age at menarche (<12, 12, 13,>13, unknown), combined parity/age at first birth (parous with first birth before age 25, parous with first birth at or after age 25, nulliparous, unknown), benign breast disease subtype (non-proliferative, proliferative without atypia, proliferative with atypia), alcohol use (none, >0 - <5, ≥ 5 g/day), and NHS cohort (NHS, NHSII); and (3) Model 3 (reduced): age (continuous), BMI (continuous), menopausal status (premenopausal, postmenopausal, unknown menopausal status), combined parity/age at first birth (parous with first birth before age 25, parous with first birth at or after age 25, nulliparous, unknown), and benign breast disease subtype (non-proliferative, proliferative without atypia, proliferative with atypia). Due to the relatively small sample size and large number of covariates, we present Model 3 as our final model. Median levels within respective categories of IGF-1 and IGFBP-3 were used for the test of trend. All the analyses were performed using SAS software (version 9.4, SAS Institute, Cary, NC). All tests of statistical significance were 2-sided.

Results

In this study of 151 women, 3 (2.0%) had non-proliferative disease, 137 (90.7%) had proliferative disease without atypia, and 11 (7.28%) had atypical hyperplasia. The average age at biopsy was 42.9 years (range 34–61). Majority of the women were premenopausal (88.7%). Age-adjusted characteristics of women in the study by IGF-1 and IGFBP-3 levels are presented in Table 1.

Age and BMI-adjusted risk estimates (Model 1) are presented in Supplementary Table 1 and the results of the fully-adjusted model (Model 2) are presented in Supplementary Table 2. In multivariate analysis (Table 2), circulating IGF-1 was not associated with the expression of any of the markers in epithelium or stroma (Table 2). Continuous measures of IGFBP-3 were not associated with the expression of CD44 or ALDH1A1 markers in epithelium or stroma. However, there was a suggestion of an inverse association between IGFBP-3 quartiles with stromal CD24 expression in normal TDLUs (β for 4th vs. 1st quartile=-0.38, 95% CI-0.69,-0.08, *p*-trend=0.06).

Discussion

In this study of 151 women, we only found significant inverse associations between circulating IGFBP-3 with stromal CD24 expression in benign breast tissues. IGF-1 is expressed in many tissues and in the breast, predominantly in the stroma of normal and malignant tissue [41]. IGF-1 signaling pathway plays a critical role in both the EMT and stem cell-related processes in normal and cancerous tissues with numerous studies showing its involvement in self renewal, stem cell surface markers, migration, and invasion and tumor initiation in lung, prostate, liver, and breast cancers [13]. Previous pooled analyses reported positive associations of circulating IGF-1 and IGFBP-3 with breast cancer [42]. Stem cell markers CD44, CD24, and ALDH1A1 in breast tumors were linked to younger age at diagnosis, unfavorable tumor characteristics, (i.e. grade, stage, and triple-negative status), metastatic spread, poor prognosis and chemotherapy resistance [16-24], with positive associations for CD44 and ALDH1A1 and inverse associations for CD24. Our current study observed a suggestive inverse association between circulating IGFBP-3 and benign

Characteristic	Below median IGF-1 n=76	Above median IGF-1 n = 75	Below median IGFBP-3 n=75	Above median IGFBP-3 n=76
Mean (SD)				
Age (years) ^a	44.62 (5.09)	41.25 (4.59)	43.92 (5.44)	41.99 (4.61)
Age at menarche (years)	12.51 (1.43)	12.79 (1.30)	12.72 (1.56)	12.76 (1.42)
Body Mass Index (kg/m ²)	26.51 (6.65)	24.47 (4.00)	25.22 (5.32)	25.55 (5.48)
Alcohol, g/day	2.67 (4.32)	2.92 (5.06)	3.28 (5.84)	2.42 (3.74)
Percentages				
Parity/age at first birth				
Nulliparous	20	15	22	14
Parous, age < 25 years	31	27	27	28
Parous, age≥25 years	47	56	49	57
Family history of breast cancer	5	5	4	9
Benign breast disease				
Non-proliferative	4	0	3	1
Proliferative without atypia	87	94	87	95
Proliferative with atypia	9	6	10	4
Menopausal status/hormone use				
Premenopausal	84	93	86	94
Postmenopausal	14	3	11	2

Table 1 Age-adjusted characteristics of Nurses' Health Study II participants, by IGF-1 levels at time of biopsy

^a Value is not age adjusted

breast stromal CD24 expression, providing indirect evidence of the effect of IGF on EMT and stem cell activity in normal TDLUs.

To our knowledge, this is the first study to date exploring the associations of IGF-1 and IGFBP-3 with breast stem cell markers. The analysis used data from NHSII, an established cohort with more than 30 years of follow-up, confirmed benign breast disease status, and comprehensive information on breast cancer risk factors. Our study has a few limitations. First, we had only a single IGF-1 and IGFBP-3 measurement. However, previous work in this cohort shows high correlations of these measures over the 3-year period among premenopausal women (interclass correlations 0.70 for IGF-1 and 0.74 for IGFBP-3), suggesting that a single measurement of IGF-1 and IGFBP-3 reliably represents their levels over a long period of time [28]. Next, we recognize that biopsy samples come from a specific area of the breast. Our previous work shows that this sampling method is able to provide strong evidence for a priori hypotheses and meaningful findings for breast tissue involution [5], identification of markers associated with breast cancer risk [12, 43, 44], and associations with known breast cancer risk factors, suggesting that this limitation has minimal impact on research findings [45]. Further, previous studies suggest that CD44(+)CD24(-/low) and ALDH1(high) expression could be used to characterize two largely non-overlapping populations of breast cancer stem cells which have epithelial-like and mesenchymal-like phenotypes, respectively [46–48]. We stained each marker separately. Hence, we were unable to determine the co-localization of these stem cell markers or assess the combination of these markers' expression on a cell-by-cell basis. Finally, we were unable to make any adjustments for circulating estrogens to account for interplay between estrogens and IGF pathway and their potential influence on stem cells as these measurements were not available in this study sample.

Conclusions

We found a suggestive association of higher circulating IGFBP-3 levels with lower stromal CD24 expression in benign breast tissue. As stem cells are characterized by higher CD44/lower CD24 expression, our findings support the hypothesis that the activation of the IGF

IGF/IGFBP-3	ð	14			CD24	-			ALD	H1A1		
	z	In Epithelium	z	In Stroma	z	In Epithelium	z	In Stroma	z	In Epithelium	z	In Stroma
IGF-1, continuous log-transformed	147	0.09 (-0.45; 0.62)	148	0.88 (- 0.30; 2.06) 1	141	0.03 (-0.16; 0.21)	143	0.22 (-0.20; 0.65)	145	0.28 (-0.49; 1.06)	143	0.36 (-1.12; 1.85)
IGF-1 quartiles												
Quartile 1	38	Ref	38	Ref 3	36	Ref	37	Ref	37	Ref	36	Ref
Quartile 2	35	-0.06 (-0.46; 0.34)	36	0.69 (0.17; 1.55) 3	33	0.02 (-0.12; 0.16)	34	0.13 (-0.18; 0.44)	36	0.35 (-0.23; 0.93)	36	0.33 (-0.79; 1.46)
Quartile 3	37	-0.08(-0.48; 0.32)	37	0.28 (0.58; 1.14) 3	37	-0.02 (-0.15; 0.12)	37	-0.01 (-0.32; 0.30)	36	0.20 (- 0.38; 0.78)	35	0.59 (-0.54; 1.72)
Quartile 4	37	-0.03 (-0.48; 0.41)	37	0.93 (0.03; 1.89)	35	0.01 (-0.14; 0.17)	35	0.25 (-0.10; 0.60)	36	0.56 (- 0.08; 1.20)	36	0.54 (-0.72; 1.79)
p-trend ^a		0.59		0.36		0.66		0.57		0.45		0.25
IGFBP-3 continuous log-transformed	147	-0.04 (-0.94; 0.85)	148	-0.41 (-2.35; 1.53) 1	141	-0.17 (-0.47; 0.13)	143	-0.60 (-1.30; 0.10)	145	0.25 (- 1.02; 1.52)	143	0.96 (1.46; 3.38)
IGFBP-3 quartiles												
Quartile 1	35	Ref	36	Ref 3	34	Ref	36	Ref	36	Ref	36	Ref
Quartile 2	37	-0.11 (-0.47; 0.26)	37	-0.15 (-0.93; 0.62) 3	35	-0.09 (-0.21; 0.04)	35	-0.27 (-0.55; 0.02)	36	-0.19 (-0.71; 0.33)	35	0.28 (-0.73; 1.28)
Quartile 3	37	0.06 (0.32; 0.44)	37	0.10 (- 0.72; 0.91) ŝ	34	-0.06 (-1.19; 0.08)	34	-0.22 (-0.52; 0.07)	35	0.05 (- 0.49; 0.59)	35	-0.23 (-1.27; 0.81)
Quartile 4	38	-0.12 (-0.51; 0.27)	38	-0.58 (-1.42; 0.26) 3	38	-0.08 (-0.21; 0.05)	38	-0.38 (-0.69; -0.08)	38	-0.18 (-0.73; 0.38)	37	0.11 (-0.98; 1.20)
p-trend ^a		0.82		0.52		0.56		0.06		0.21		0.77

posumenopausat, anknowny, and bengin breast disease subtype (non-promerauve) ^a *p*-trend test performed using median values within each quartile

Abbreviations

BBD	Benign breast disease
BMI	Body mass index
CI	Confidence interval
EMT	Epithelial-to-mesenchymal transition
IGF	Insulin-like growth factor
IGFBP-3	Insulin-like growth factor-binding protein-3
NHS	Nurses' Health Study
TDLU	Terminal duct-lobular units
TMA	Tissue microarray

Supplementary Information

The online version contains supplementary material available at https://doi. org/10.1186/s13058-025-02002-z.

Additional file 1.

Acknowledgements

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.

Author contributions

LY and RT conceived of and designed the study, directed statistical analyses, interpreted results, substantially revised initial drafts of the paper and provided final review and approval. LY performed statistical analyses. YH and BS assessed IHC results. LY wrote the first draft of the manuscript which was revised with contribution from YH, GB, DM, MM, BR, KP, MR, and RT. All authors read and approved the final manuscript. LY supervised the overall study progress. All authors read and approved the final manuscript.

Funding

This work was supported by the National Cancer Institute at the National Institutes of Health [CA240341 to L.Y., CA131332, CA175080, P01 CA087969 to R.M.T., UM1 CA186107 and, to M.S., U01 CA176726 to W.W], Avon Foundation for Women, Susan G. Komen for the Cure[®], and Breast Cancer Research Foundation.

Availability of data and materials

The data that support the findings of this study are available from the Nurses' Health Studies, however they are not publicly available. Investigators interested in using the data can request access, and feasibility will be discussed at an investigators' meeting. Limits are not placed on scientific questions or methods, and there is no requirement for co-authorship. Additional data sharing information and policy details can be accessed at http://www.nursesheal thstudy.org/researchers.

Declarations

Ethical approval and consent to participate

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. Consent was obtained or implied by return of questionnaires.

Consent for publication

Not applicable.

Competing interests

All funding sources have been acknowledged. There are no other competing interests.

Author details

¹ Department of Epidemiology, College of Public Health and Health Professions and College of Medicine, University of Florida, 2004 Mowry Rd., Gainesville, FL 32610, USA. ² Department of Pathology, Beth Israel Deaconess Medical Center, Harvard Medical School, Boston, MA, USA. ³Channing Division of Network Medicine, Department of Medicine, Brigham and Women's Hospital and Harvard Medical School, Boston, MA, USA. ⁴Dana-Farber Cancer Institute, Brigham and Women's Hospital, and Harvard Medical School, Boston, MA, USA. ⁵Department of Population Health Sciences, Weill Cornell Medicine, New York, NY, USA.

Received: 17 December 2024 Accepted: 12 March 2025 Published online: 07 April 2025

References

- Cobaleda C, Cruz JJ, Gonzalez-Sarmiento R, Sanchez-Garcia I, Perez-Losada J. The emerging picture of human breast cancer as a stem cellbased disease. Stem Cell Rev. 2008;4(2):67–79.
- Smalley M, Ashworth A. Stem cells and breast cancer: a field in transit. Nat Rev Cancer. 2003;3(11):832–44.
- Chang C-C. Recent translational research: stem cells as the roots of breast cancer. Breast Cancer Res. 2006;8(1):103–103.
- 4. Ginestier C, Wicha MS. Mammary stem cell number as a determinate of breast cancer risk. Breast Cancer Res. 2007;9(4):109.
- Rice MS, Tamimi RM, Connolly JL, Collins LC, Shen D, Pollak MN, Rosner B, Hankinson SE, Tworoger SS. Insulin-like growth factor-1, insulin-like growth factor binding protein-3 and lobule type in the Nurses' Health Study II. Breast Cancer Res. 2012;14(2):1–7.
- Lagiou P, Samoli E, Lagiou A, Zourna P, Barbouni A, Georgila C, Tsikkinis A, Vassilarou D, Minaki P, Sfikas C, et al. A comparison of hormonal profiles between breast cancer and benign breast disease: a case–control study. Ann Oncol. 2013. https://doi.org/10.1093/annonc/mdt207.
- Samoli E, Lagiou A, Zourna P, Barbouni A, Georgila C, Tsikkinis A, Vassilarou D, Minaki P, Sfikas C, Spanos E, et al. Expression of estrogen receptors in non-malignant mammary tissue modifies the association between insulin-like growth factor 1 and breast cancer risk. Ann Oncol: Offic J Eur Soc Med Oncol ESMO. 2015;26(4):793–7.
- Su X, Colditz GA, Willett WC, Collins LC, Schnitt SJ, Connolly JL, Pollak MN, Rosner B, Tamimi RM. Genetic variation and circulating levels of IGF-I and IGFBP-3 in relation to risk of proliferative benign breast disease. Int J Cancer J Int Cancer. 2010;126(1):180–90.
- Christopoulos PF, Msaouel P, Koutsilieris M. The role of the insulin-like growth factor-1 system in breast cancer. Mol Cancer. 2015;14:43.
- Byrne C, Colditz GA, Willett WC, Speizer FE, Pollak M, Hankinson SE. Plasma insulin-like growth factor (IGF) I, IGF-binding protein 3, and mammographic density. Can Res. 2000;60(14):3744–8.
- Schernhammer ES, Tworoger SS, Eliassen AH, Missmer SA, Holly JM, Pollak MN, Hankinson SE. Body shape throughout life and correlations with IGFs and GH. Endocr Relat Cancer. 2007;14(3):721–32.
- Tamimi RM, Colditz GA, Wang Y, Collins LC, Hu R, Rosner B, Irie HY, Connolly JL, Schnitt SJ. Expression of IGF1R in normal breast tissue and subsequent risk of breast cancer. Breast Cancer Res Treat. 2011;128(1):243–50.
- Farabaugh SM, Boone DN, Lee AV. Role of IGF1R in breast cancer subtypes, stemness, and lineage differentiation. Front Endocrinol. 2015;6:59.
- 14. Malaguarnera R, Belfiore A. The emerging role of insulin and insulin-like growth factor signaling in cancer stem cells. Front Endocrinol. 2014;5:10.
- Anwar T, Kleer CG. Tissue-based identification of stem cells and epithelial-to-mesenchymal transition in breast cancer. Hum Pathol. 2013;44(8):1457–64.
- Giatromanolaki A, Sivridis E, Fiska A, Koukourakis M. The CD44+/CD24– phenotype relates to 'triple-negative' state and unfavorable prognosis in breast cancer patients. Med Oncol. 2010;28:745.

- Abraham BK, Fritz P, Van der Kuip H, Buck M, Szabo Z, Athelogou M, Brauch H. Evaluation of CD44+/CD24-/low cells in breast cancer and relevance for distant metastasis. AACR Meet Abstracts. 2005;2005(1):481.
- Neumeister V, Agarwal S, Bordeaux J, Camp RL, Rimm DL. In situ identification of putative cancer stem cells by multiplexing ALDH1, CD44, and cytokeratin identifies breast cancer patients with poor prognosis. Am J Pathol. 2010;176(5):2131–8.
- Lee JH, Kim SH, Lee ES, Kim YS. CD24 overexpression in cancer development and progression: a meta-analysis. Oncol Rep. 2009;22(5):1149–56.
- Eden JA. Human breast cancer stem cells and sex hormones–a narrative review. Menopause. 2010;17(4):801–10.
- Ginestier C, Hur MH, Charafe-Jauffret E, Monville F, Dutcher J, Brown M, Jacquemier J, Viens P, Kleer CG, Liu S, et al. ALDH1 is a marker of normal and malignant human mammary stem cells and a predictor of poor clinical outcome. Cell Stem Cell. 2007;1 (5):555–67.
- 22. Liu Y, Lv DL, Duan JJ, Xu SL, Zhang JF, Yang XJ, Zhang X, Cui YH, Bian XW, Yu SC. ALDH1A1 expression correlates with clinicopathologic features and poor prognosis of breast cancer patients: a systematic review and meta-analysis. BMC Cancer. 2014;14:444.
- Khoury T, Ademuyiwa FO, Chandraseekhar R, Jabbour M, DeLeo A, Ferrone S, Wang Y, Wang X. Aldehyde dehydrogenase 1A1 expression in breast cancer is associated with stage, triple negativity, and outcome to neoadjuvant chemotherapy. Mod Pathol. 2012;25(3):388–97.
- Yao J, Jin Q. Wang X-d, Zhu H-j, Ni Q-c: Aldehyde dehydrogenase 1 expression is correlated with poor prognosis in breast cancer. Medicine. 2017;96(25):e7171.
- 25. Tamimi RM, Byrne C, Colditz GA, Hankinson SE. Endogenous hormone levels, mammographic density, and subsequent risk of breast cancer in postmenopausal women. J Natl Cancer Inst. 2007;99(15):1178–87.
- Collins LC, Baer HJ, Tamimi RM, Connolly JL, Colditz GA, Schnitt SJ. The influence of family history on breast cancer risk in women with biopsyconfirmed benign breast disease: results from the Nurses' Health Study. Cancer. 2006;107(6):1240–7.
- 27. Tworoger SS, Sluss P, Hankinson SE. Association between plasma prolactin concentrations and risk of breast cancer among predominately premenopausal women. Cancer Res. 2006;66(4):2476–82.
- Rice MS, Tamimi RM, Connolly JL, Collins LC, Shen D, Pollak MN, Rosner B, Hankinson SE, Tworoger SS. Insulin-like growth factor-1, insulin-like growth factor binding protein-3 and lobule type in the Nurses' Health Study II. Breast Cancer Res. 2012;14(2):R44.
- Jacobs TW, Byrne C, Colditz G, Connolly JL, Schnitt SJ. Pathologic features of breast cancers in women with previous benign breast disease. Am J Clin Pathol. 2001;115(3):362–9.
- Tamimi RM, Byrne C, Baer HJ, Rosner B, Schnitt SJ, Connolly JL, Colditz GA. Benign breast disease, recent alcohol consumption, and risk of breast cancer: a nested case–control study. Breast Cancer Res. 2005;7(4):R555.
- Page DL, Dupont WD, Rogers LW, Rados MS. Atypical hyperplastic lesions of the female breast: a long-term follow-up study. Cancer. 1985;55(11):2698–708.
- Collins LC, Wang YH, Connolly JL, Baer HJ, Hu R, Schnitt SJ, Colditz GA, Tamimi RM. potential role of tissue microarrays for the study of biomarker expression in benign breast disease and normal breast tissue. Appl Immunohisto M M. 2009;17(5):438–41.
- Yaghjyan L, Heng YJ, Baker GM, Bret-Mounet V, Murthy D, Mahoney MB, Mu Y, Rosner B, Tamimi RM. Reliability of CD44, CD24, and ALDH1A1 immunohistochemical staining: pathologist assessment compared to quantitative image analysis. Front Med. 2022. https://doi.org/10.3389/ fmed.2022.1040061.
- Yaghjyan L, Stoll E, Ghosh K, Scott CG, Jensen MR, Brandt KR, Visscher D, Vachon CM. Tissue-based associations of mammographic breast density with breast stem cell markers. Breast Cancer Res BCR. 2017;19(1):100–100.
- Yaghjyan L, Esnakula AK, Scott CG, Wijayabahu AT, Jensen MR, Vachon CM. Associations of mammographic breast density with breast stem cell marker-defined breast cancer subtypes. Cancer Causes Control. 2019;30(10):1103–11.
- Bankhead P, Loughrey MB, Fernández JA, Dombrowski Y, McArt DG, Dunne PD, McQuaid S, Gray RT, Murray LJ, Coleman HG, et al. QuPath: Open source software for digital pathology image analysis. Sci Rep. 2017;7(1):16878.
- Baker GM, Bret-Mounet VC, Wang T, Veta M, Zheng H, Collins LC, Eliassen AH, Tamimi RM, Heng YJ. Immunohistochemistry scoring of breast tumor

tissue microarrays: a comparison study across three software applications. J Pathol Inform. 2022;13:100118.

- Willett W, Stampfer MJ, Bain C, Lipnick R, Speizer FE, Rosner B, Cramer D, Hennekens CH. Cigarette smoking, relative weight, and menopause. Am J Epidemiol. 1983;117(6):651–8.
- Stampfer MJ, Willett WC, Colditz GA, Rosner B, Speizer FE, Hennekens CH. A prospective study of postmenopausal Estrogen therapy and coronary heart disease. N Engl J Med. 1985;313(17):1044–9.
- 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. 2008;167(6):653–66.
- Macias H, Hinck L. Mammary gland development. Wiley Interdiscip Rev Dev Biol. 2012;1(4):533–57.
- Chang CI, Low HP, Qiu L, Strohsnitter WC, Hsieh CC. Prenatal modulation of breast density and breast stem cells by insulin-like growth factor-1. Am J Stem Cells. 2012;1(3):239–52.
- Huh SJ, Oh H, Peterson MA, Almendro V, Hu R, Bowden M, Lis RL, Cotter MB, Loda M, Barry WT, et al. The proliferative activity of mammary epithelial cells in normal tissue predicts breast cancer risk in premenopausal women. Can Res. 2016;76(7):1926–34.
- 44. Oh H, Eliassen AH, Wang M, Smith-Warner SA, Beck AH, Schnitt SJ, Collins LC, Connolly JL, Montaser-Kouhsari L, Polyak K, et al. Expression of estrogen receptor, progesterone receptor, and Ki67 in normal breast tissue in relation to subsequent risk of breast cancer. NPJ Breast Cancer. 2016;2:16032.
- 45. Oh H, Eliassen AH, Beck AH, Rosner B, Schnitt SJ, Collins LC, Connolly JL, Montaser-Kouhsari L, Willett WC, Tamimi RM. Breast cancer risk factors in relation to estrogen receptor, progesterone receptor, insulin-like growth factor-1 receptor, and Ki67 expression in normal breast tissue. NPJ Breast Cancer. 2017;3(1):39.
- Colacino JA, Azizi E, Brooks MD, Harouaka R, Fouladdel S, McDermott SP, Lee M, Hill D, Madden J, Boerner J, et al. Heterogeneity of human breast stem and progenitor cells as revealed by transcriptional profiling. Stem Cell Reports. 2018;10(5):1596–609.
- Ricardo S, Vieira AF, Gerhard R, Leitão D, Pinto R, Cameselle-Teijeiro JF, Milanezi F, Schmitt F, Paredes J. Breast cancer stem cell markers CD44, CD24 and ALDH1: expression distribution within intrinsic molecular subtype. J Clin Pathol. 2011;64(11):937–46.
- Escudero Mendez L, Srinivasan M, Hamouda RK, Ambedkar B, Arzoun H, Sahib I, Fondeur J, Mohammed L. Evaluation of CD44+/CD24- and aldehyde dehydrogenase enzyme markers in cancer stem cells as prognostic indicators for triple-negative breast cancer. Cureus. 2022;14(8):e28056.

Publisher's Note

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