Review

Cancer Epidemiology, Biomarkers & Prevention

The Premenopausal Breast Cancer Collaboration: A Pooling Project of Studies Participating in the National Cancer Institute Cohort Consortium

Hazel B. Nichols[1], Minouk J. Schoemaker[2], Lauren B. Wright[2], Craig McGowan[1], Mark N. Brook[2], Kathleen M. McClain[1], Michael E. Jones[2], Hans-Olov Adami[3], Claudia Agnoli[4], Laura Baglietto[5], Leslie Bernstein[6], Kimberly A. Bertrand[7], William J. Blot[8], Marie-Christine Boutron-Ruault[5], Lesley Butler[9], Yu Chen[10], Michele M. Doody[11], Laure Dossus[12], A. Heather Eliassen[13], Graham G. Giles[14], Inger T. Gram[15], Susan E. Hankinson[16], Judy Hoffman-Bolton[17], Rudolf Kaaks[18], Timothy J. Key[19], Victoria A. Kirsh[20], Cari M. Kitahara[11], Woon-Puay Koh[21], Susanna C. Larsson[22], Eiliv Lund[23], Huiyan Ma[6], Melissa A. Merritt[24], Roger L. Milne[14], Carmen Navarro[25,26], Kim Overvad[27], Kotaro Ozasa[28], Julie R. Palmer[7], Petra H. Peeters[29], Elio Riboli[24], Thomas E. Rohan[30], Atsuko Sadakane[28], Malin Sund[31], Rulla M. Tamimi[13], Antonia Trichopoulou[32], Lars Vatten[33], Kala Visvanathan[17,34], Elisabete Weiderpass[35,36,37,38], Walter C. Willett[39], Alicja Wolk[22], Anne Zeleniuch-Jacquotte[10], Wei Zheng[8], Dale P. Sandler[40], and Anthony J. Swerdlow[2,41]

Abstract

Breast cancer is a leading cancer diagnosis among premenopausal women around the world. Unlike rates in postmenopausal women, incidence rates of advanced breast cancer have increased in recent decades for premenopausal women. Progress in identifying contributors to breast cancer risk among premenopausal women has been constrained by the limited numbers of premenopausal breast cancer cases in individual studies and resulting low statistical power to subcategorize exposures or to study specific subtypes. The Premenopausal Breast Cancer Collaborative Group was established to facilitate cohort-based analyses of risk factors for premenopausal breast cancer by pooling individual-level data from studies participating in the United States National Cancer Institute Cohort Consortium. This article describes the Group, including the rationale for its initial aims related to pregnancy, obesity, and physical activity. We also describe the 20 cohort studies with data submitted to the Group by June 2016. The infrastructure developed for this work can be leveraged to support additional investigations. *Cancer Epidemiol Biomarkers Prev; 26(9); 1360–9. ©2017 AACR.*

[1]University of North Carolina, Gillings School of Global Public Health, Chapel Hill, North Carolina. [2]Division of Genetics and Epidemiology, The Institute of Cancer Research, Sutton, United Kingdom. [3]Karolinska Institutet, MEB, University of Oslo Institute of Health and Society, Sweden, Norway. [4]Fondazione IRCCS Istituto Nazionale dei Tumori, Italy, France. [5]Centre for Research in Epidemiology and Population Health (CESP), France. [6]Beckman Research Institute of City of Hope, Monrovia, California. [7]Slone Epidemiology Center at Boston University, Boston, Massachusetts. [8]Division of Epidemiology, Department of Medicine, Vanderbilt-Ingram Cancer Center, Vanderbilt University School of Medicine, Nashville, Tennessee. [9]University of Pittsburgh Graduate School of Public Health, University of Pittsburgh Cancer Institute, Pittsburgh, Pennsylvania. [10]NYU School of Medicine, New York, New York. [11]Radiation Epidemiology Branch, Division of Cancer Epidemiology and Genetics, National Cancer Institute, NIH, Bethesda, Maryland. [12]International Agency for Research on Cancer, France. [13]Channing Division of Network Medicine, Brigham and Women's Hospital, Harvard T.H. Chan School of Public Health and Harvard Medical School, Boston, Massachusetts. [14]Cancer Council Victoria, University of Melbourne, Melbourne, Australia. [15]University of Tromsø, The Arctic University of Norway, Tromsø, Norway. [16]Department of Biostatistics and Epidemiology, School of Public Health and Health Sciences, University of Massachusetts, Boston, Massachusetts. [17]Johns Hopkins Bloomberg School of Public Health, Baltimore, Maryland. [18]Division of Cancer Epidemiology, DKFZ, Heidelberg, Germany. [19]University of Oxford, Oxford, United Kingdom. [20]Dalla Lana School of Public Health, University of Toronto, Toronto, Ontario, Canada. [21]Duke-NUS Medical School, Singapore. [22]Karolinska Institutet, Institute of Environmental Medicine, Stockholm, Sweden. [23]UiT (University of Tromsø), Tromsø, Norway. [24]School of Public Health, Imperial College London, United Kingdom. [25]Department of Epidemiology, Murcia Regional Health Council, IMIB-Arrixaca, Murcia, Spain. [26]CIBER Epidemiología y Salud Pública (CIBERESP); Department of Health and Social Sciences, Universidad de Murcia, Murcia, Spain. [27]Department of Public Health, Section for Epidemiology, Aarhus University, Aarhus, Denmark. [28]Radiation Effects Research Foundation, Hiroshima, Japan. [29]University Medical Center, Utrecht, the Netherlands. [30]Albert Einstein College of Medicine, Bronx, New York. [31]Umeå University, Umeå, Sweden. [32]Hellenic Health Foundation, Greece. [33]Norwegian University of Science and Technology, Trondheim, Norway. [34]Johns Hopkins School of Medicine, Baltimore, Maryland. [35]Karolinska Institutet, Department of Epidemiology and Biostatistics, Oslo, Norway. [36]Department of Research, Head, Group of Etiological Cancer Research, Cancer Registry of Norway, Institute of Population-Based Cancer Research, Oslo, Norway. [37]Genetic Epidemiology Group, Folkhälsan Research Center, Helsinki, Finland. [38]Department of Community Medicine, University of Tromsø, The Arctic University of Norway, Tromsø, Norway. [39]Department of Nutrition and Epidemiology, Harvard T.H. Chan School of Public Health, Boston, Massachusetts. [40]Epidemiology Branch, National Institute of Environmental Health Sciences, National Institutes of Health, Durham, North Carolina. [41]Division of Breast Cancer Research, The Institute of Cancer Research, London, United Kingdom.

Introduction

Breast cancer is the most common cancer diagnosed among women worldwide, with an estimated 1.67 million cases diagnosed in 2012, accounting for a quarter of all new cancers in women. Breast cancer is also the most common cancer diagnosed among women ages 15 to 39 years worldwide (1). Furthermore, breast cancer among premenopausal women often presents at more advanced stages (2) and, at the youngest ages, has less favorable prognosis (3) than among postmenopausal women.

Identifying contributors to breast cancer risk in younger women is critical to prevention. In the United States, incidence rates of advanced breast cancer have increased among premenopausal women in recent decades, whereas they have consistently decreased among women ages 60 and older during the same period (4). Accumulating evidence supports etiologic heterogeneity between pre- and postmenopausal breast cancer. Several lifestyle factors, including childbirth (5), obesity (6), and cigarette smoking (7), have been reported to have differential associations with breast cancer risk before and after menopause. Breast cancer subtypes, including those defined by gene expression (8), or clinical markers including estrogen receptor (ER), progesterone receptor (PR), or HER2/*neu* oncogene expression, have emerged as critical considerations for risk factor associations and are differentially distributed by menopausal status (9). Investigations of breast cancer etiologic heterogeneity require large sample sizes to have sufficient statistical power to account jointly for menopausal status and tumor subtype.

The Premenopausal Breast Cancer Collaborative Group (referred to hereafter as the Collaborative Group) was established to facilitate cohort-based analyses of risk factors for premenopausal breast cancer, both overall and according to tumor characteristics. This article describes the formation of the Collaborative Group, the methods used for ongoing efforts, and provides the rationale for initial analyses related to pregnancy, obesity, and physical activity. The infrastructure developed to address these questions can support future investigations of additional potential risk factors.

Collaborative Group Studies

The National Cancer Institute (NCI) Cohort Consortium was formed to address the need for large-scale collaborations to pool data in cohort studies of cancer and hence to quicken the pace of research (http://epi.grants.cancer.gov/Consortia/cohort.html). The Collaborative Group was initiated within the Cohort Consortium in 2013 by investigators at The Institute of Cancer Research (ICR) in London and the National Institute of Environmental Health Sciences (NIEHS). The ICR and the NIEHS serve as the Data Coordinating Centers.

Eligibility

Prospective cohorts in the Cohort Consortium with at least 100 female breast cancers diagnosed during follow-up before age 55 and data collection at two or more time points (baseline and at least one follow-up, to allow for exposure information and menopausal status to be updated) are eligible to participate.

Participating cohorts

This report describes the 20 cohort studies (counting the European Prospective Investigation into Cancer and Nutrition, which has many cohorts within it, as a single cohort; refs. 6, 10–28) with data submitted to the Collaborative Group as of June 2016. Participating cohorts are shown in Table 1 and span North America, Europe, Asia, and Australia. The number of female participants from these cohorts ages <55 at enrollment ranges from 5,671 (Campaign against Cancer and Heart Disease) to 117,733 (Nurses' Health Study cohort). The cohorts were initiated as early as 1950 (the Radiation Effects Research Foundation Life Span Study) or as recently as 2003 (Generations and Sister Study cohorts). All cohorts have conducted more than one round of data collection; however, follow-up data are not yet fully available for three cohorts. The number of follow-up rounds for which data have been submitted as of June 1, 2016, ranged from 1 to 16 across cohorts.

Breast cancer ascertainment

To date, data have been received for 1,030,761 women, and include 21,766 incidents invasive or *in situ* breast cancers diagnosed after study enrollment and before age 55 years (Table 2). Across studies, cancer diagnoses are identified by linkage with city/state/provincial/regional (10, 12, 13, 23, 28–31) or national (11, 12, 14, 24, 26, 32, 33) population-based cancer registries, and/or through self-report followed by medical record review (6, 10, 11, 14, 15, 25, 34, 35). All participating studies established case ascertainment procedures and published findings related to incident breast cancer risk prior to joining the Collaborative Group.

Data exchange and harmonization

After approval by the NCI Cohort Consortium executive committee, the aims of the proposed collaboration were circulated to all Consortium members in 2013. Key exposure, covariate, and outcome information necessary to address the initial analyses and potential confounding or effect modification was identified by the Coordinating Centers. Complete capture of all information across exposures is not required for participation.

After confirming eligibility, a data request template is sent to cohorts that wish to participate. Requested exposure data include: age/year of cohort entry, length of follow-up, demographic characteristics (age, race/ethnicity, education, socioeconomic status), lifestyle factors (physical activity, anthropometric characteristics, alcohol intake, smoking information, mammography use), reproductive history (menarche, menstrual cycle characteristics, gravidity, parity, pregnancy complications, infertility, breastfeeding, hormonal medications, menopausal status), benign breast disease, and family history of breast cancer (Supplementary Table S1). Most characteristics are requested at enrollment and each follow-up, as available. Breast cancer information includes age at diagnosis, stage, grade, histology, and expression

Note: Supplementary data for this article are available at Cancer Epidemiology, Biomarkers & Prevention Online (http://cebp.aacrjournals.org/).

H.B. Nichols and M.J. Schoemaker are co-first authors of this article.

D.P. Sandler and A.J. Swerdlow are co-senior authors of this article.

Corresponding Author: Hazel B. Nichols, University of North Carolina Gillings School of Global Public Health, 135 Dauer Drive, Chapel Hill, NC 27599. Phone: 919-966-7456; Fax: 919-962-2089; E-mail: hazel.nichols@unc.edu

doi: 10.1158/1055-9965.EPI-17-0246

Table 1. Characteristics of women younger than 55 years in cohorts included in the Premenopausal Breast Cancer Collaborative Group

Cohort	Location	Ages at enrollment; mean (SD), range	Calendar years of enrollment	Baseline data collection methods	*N* of data collection rounds[a]	Breast cancer cases (*N*)	Breast cancer ascertainment sources	Cohort size (women <55 years)	*N* years of follow-up, mean (SD), range (<55 years)
Black Women's Health Study (10)	United States	37.1 (8.6) 20–54	1995	Mailed questionnaire	9	1,299	Self-report and state registry	52,543	12.6 (5.6) 0–18.6
California Teachers Cohort (28)	United States	40.4 (7.4) 22–54	1995–1998	Mailed questionnaire	4	1,185	State registry	47,516	11.6 (5.0) 0.0–17.2
Campaign against Cancer and Heart Disease (CLUE II; ref. 13)	United States	39.6 (9.6) 18–54	1989	Administered questionnaire	6	131	State registry	5,671	10.8 (5.4) 0.3–26.0
Canadian Study of Diet, Lifestyle, and Health (12)[b]	Canada	44.1 (6.9) 23–54	1991–1999	Mailed questionnaire	1	377	Provincial and national registry	1,589	8.1 (4.7) 0–18.6
European Prospective Investigation into Cancer and Nutrition (14)[c]	Europe	44.2 (8.1) 19–54	1991–2000	Self-reported/ administered questionnaires	1	2,122	Self-report and national/regional registries	150,291	7.5 (4.2) 0–16.6
Etude Epidémiologique auprès de femmes de la Mutuelle Générale de l'Education Nationale (E3N; ref. 15)	France	46.5 (4.2) 38–54	1989–1991	Mailed questionnaire	8	1,908	Self-report	72,748	8.1 (4.2) 0–16.4
Generations Study (11)	United Kingdom	39.8 (9.5) 16–54	2003–2012	Mailed questionnaire	2	719	Self-report and national registry	72,058	5.4 (1.7) 0–9.7
Helseundersøkelsen i Nord-Trøndelag (HUNT2; ref. 26)	Norway	38.9 (9.7) 20–54	1995–1997	Administered questionnaire	1	209	National cancer registry	20,974	10.2 (4.1) 0.3–14.0
Melbourne Collaborative Cohort Study (16)	Australia	47.5 (4.4) 31–54	1990–1994	Administered questionnaire	3	227	State registry	12,029	7.3 (4.4) 0–20.1
New York University Women's Health Study (19, 20)	United States	45.2 (5.5) 31–54	1984–1991	Self-administered questionnaire	6	371	Self-report and state registry	8,757	9.5 (5.5) 0–23.5
Norwegian Women and Cancer Study (105)	Norway	45.7 (6.0) 31–54	1991–2007	Mailed questionnaire	3	2,124	National registry	117,633	9.0 (5.8) 0.3–20.5
Nurses' Health Study (17)	United States	42.6 (7.1) 29–54	1976–1978	Mailed questionnaire	16	2,743	Self-report	117,730	12.2 (7.0) 0.1–25.5
Nurses' Health Study II (18)	United States	34.8 (4.7) 24–44	1989–1990	Mailed questionnaire	12	3,765	Self-report	116,415	18.7 (3.7) 0.1–23.7
Radiation Effects Research Foundation Life Span Study (21)	Japan	41.3 (8.5) 18–54	1963–1993	Administered or mailed questionnaire	6	130	City registry	18,420	13.5 (8.5) 0.1–36.7
Singapore Chinese Health Study (22)	Singapore	49.6 (3.0) 43–54	1993–1998	Administered questionnaire	2	134	National cancer registry	16,056	5.3 (3.0) 0.3–11.5
Sister Study (6)	United States	47.9 (4.9) 35–54	2003–2009	Telephone and written questionnaire	3	679	Self-report	24,044	4.7 (2.5) 0.1–10.6
Southern Community Cohort Study (23)	United States	47.3 (4.2) 40–54	2002–2009	Administered questionnaire	2	233	State registry	30,289	5.1 (2.4) 0.1–13.3

(Continued on the following page)

Table 1. Characteristics of women younger than 55 years in cohorts included in the Premenopausal Breast Cancer Collaborative Group (Cont'd)

Cohort	Location	Ages at enrollment; mean (SD), range	Calendar years of enrollment	Baseline data collection methods	*N* of data collection rounds[a]	Breast cancer cases (*N*)	Breast cancer ascertainment sources	Cohort size (women <55 years)	*N* years of follow-up, mean (SD), range (<55 years)
Sweden Women's Lifestyle and Health Study (27)	Sweden	39.7 (5.8) 29–49	1991–1992	Mailed questionnaire	2	1,192	National registry	49,010	14.4 (5.3) 0.1–21.1
Swedish Mammography Cohort (24)	Sweden	46.6 (4.3) 38–54	1987–1990	Mailed questionnaire	2	649	National registry	34,126	8.3 (4.3) 0–16.6
United States Radiologic Technologist Cohort (25)	United States	36.8 (7.3) 22–54	1983–1998	Mailed questionnaire	3	1,570	Self-report	62,862	14.5 (5.6) 0–22.8

[a]Contributed as of June 2016; includes baseline and each follow-up.
[b]The Canadian Study of Diet, Lifestyle, and Health is the only case-cohort study. The cohort size ($N = 1,589$) represents the subcohort only.
[c]The European Prospective Study into Cancer and Nutrition (EPIC) dataset does not include the French or Norwegian EPIC sites which contributed from the Etude Epidémiologique auprès de femmes de la Mutuelle Générale de l'Education Nationale and Norwegian Women and Cancer Study directly.

of ER, PR, HER2, CK5/6, or EGFR. Participating studies are asked where possible to recode their own data to fit the data request template to minimize the potential for error in the recoding or understanding of variables in their original form. However, if this is not possible due to programming support constraints or other reasons, data are sent to the Coordinating Centers in their original form with a study-specific contact person identified to address questions from Coordinating Center programmers who reformat the information to fit the standard definitions in the data request template.

After data transfer agreements are signed between each individual study and the Coordinating Centers, completed datasets were transferred to the coordinating centers using secure file transfer protocols. Each cohort submits their data to one of the two Coordinating Centers who are responsible for data transfer and harmonization procedures. By having two data coordinating sites, one located in the United States and the other in the United Kingdom, we are able to minimize time zone differences to facilitate rapid communication, and accommodate studies that are only able to send data to certain locations because of country-specific information governance requirements.

Data harmonization procedures are standardized across Coordinating Centers. Quality control checks are run on each dataset to identify (i) potential data inconsistencies for each questionnaire round (e.g., nulliparous women reporting more than zero births), (ii) inconsistencies between questionnaire rounds (e.g., number of births at follow-up being lower than at baseline questionnaire), and (iii) implausible values. Data checking procedures are automated with a shared program that was run at each Coordinating Center with standardized output. Each cohort is contacted regarding any issues that were identified, and clarifications or updates are incorporated into the study-specific dataset. Where issues could not be resolved, pre-established recoding rules are applied to the data. When study-specific variables cannot be recoded to meet the data template formats (e.g., age at exposure was collected in categories but a continuous variable was requested), differences are documented and original data are retained for potential future use. Once the datasets are recoded to the standardized formats, data are merged to create a pooled dataset containing values from all cohorts.

Defining menopausal status

A primary issue for the Collaborative Group analyses is the definition of menopausal status during follow-up and at diagnosis. Menopausal status was ascertained by cohorts at each follow-up round for which it was available. In addition, we request at least one follow-up round after age 55 or breast cancer diagnosis (if available) to allow menopausal status to be defined retrospectively. In analyses conducted by menopausal status we will explore different lag periods to determine patterns for "premenopausal" or "perimenopausal" breast cancer, as menopause can be a gradual transition.

Statistical approach

Two statistical approaches are being used to analyze the data. We first examine study-specific estimates and a pooled estimate across studies using a random-effects model that weights estimates by the inverse of the study-specific variance (36–38). An advantage of this approach is that each study-specific estimate can be derived based on its own available covariates. Cochran's *Q* statistic is used to examine statistical heterogeneity between

Table 2. Breast cancer characteristics among women younger than 55 years across the Premenopausal Breast Cancer Collaborative Group

Characteristic	Combined (*N*)	Total *N* studies with data available[a]
Total breast cancers diagnosed	21,766	20 (all)
Age at diagnosis (y)		20 (all)
<30	32	
30–39	1,245	
40–44	3,340	
45–49	7,053	
50–54	10,096	
Extent of disease		20
In situ	3,651	
Invasive	17,357	
Missing	758	
Estrogen receptor status		16
Positive	9,583	
Negative	3,182	
Borderline	52	
Missing	8,949	
Progesterone receptor status		16
Positive	7,919	
Negative	3,939	
Borderline	95	
Missing	9,813	
HER2/*neu* overexpression		11
Positive	1,093	
Negative	4,808	
Borderline	29	
Missing	15,836	

[a]Contributed as of June 2016.

studies by comparing a weighted measure of difference between individual study estimates and the pooled estimate (39, 40). We calculate the I^2 statistic to examine the proportion of variance that is due to between-study heterogeneity rather than chance (41). Potential sources of heterogeneity are investigated.

Maximum flexibility for confounder adjustment and assessment of effect modification can be achieved by pooling individual-level data across cohorts. If homogeneity assumptions are not violated, we pool data into a single dataset to conduct aggregate analyses stratified by study and adjusted for potential confounders that are available in all included studies.

In both approaches, Cox regression models are used to calculate HRs and 95% confidence intervals (CI) for breast cancer (42). Regression models are constructed with age as the time scale such that person-time is accrued from age at cohort entry until breast cancer diagnosis, age at last follow-up, or other exit age, whichever occurred first. Follow-up time is stratified by time-updated exposures obtained from follow-up questionnaires, as appropriate. We test the proportional hazards assumption for exposures of interest, and in case of time-varying associations, for example, an interaction between attained age and the risk factor of interest, we investigate the addition of time-varying covariates in the model. In pooled analyses, potential variation in the association between exposures and breast cancer risk according to tumor subtype are assessed using Cox proportional hazards regression accounting for alternative tumor subtypes as competing risks (43, 44).

Rationale for Initial Aims

Pregnancy

A "dual effect" of pregnancy on breast cancer risk has been used to describe the short-term increase in breast cancer risk observed after childbirth followed by a long-term protective effect of parity. This pattern has been reported in epidemiological studies nested within European population registries (45–49) and in other case–control (50–55) and cohort (56) studies. Observational studies have reported 1.25- to 3-fold increases in breast cancer risk for up to 10 years after the last birth (5, 57). The magnitude of the pregnancy-related increase in breast cancer risk varies across studies, and may be influenced by maternal, pregnancy, or post-partum characteristics. Although a period of increased breast cancer risk after childbirth has been reported across several studies, it remains unclear whether this observation is different for, or limited to, specific groups defined by age (5, 50, 51), parity (45, 52, 53), oral contraceptive use (58), breastfeeding practices, family history of breast cancer (48, 59), or varies by breast cancer subtype (55, 56, 60) or other tumor characteristics (61, 62).

Women who have a first birth at an older age may have the greatest initial increase in breast cancer risk, and the longest interval until a protective effect appears (5, 49, 54, 63). Over the last 50 years, more women have postponed childbirth to older ages (5); this trend may have contributed to the increasing advanced-stage breast cancer rates among reproductive-age women. Pregnancy may also have opposite effects on risks of different breast cancer subtypes. For example, without considering menopausal status or subtype, parity reduces overall breast cancer risk by approximately 30% (64). However, parous women have a 50% to 90% increased risk of basal-like or ER^-/PR^- breast cancer overall (56, 65, 66). Associations for pre- and postmenopausal breast cancer combined often reflect patterns among the majority postmenopausal breast cancer cases. Our study will be well positioned to examine potential variation in the association between recent pregnancy and breast cancer subtype among premenopausal women. Others have proposed that pregnancy-related increases in breast cancer risk may also be affected by the relatively greater influence of genetic predisposition at younger versus older ages at diagnosis (48). In support of this hypothesis, at least two studies have shown stronger associations with recent birth and breast cancer risk among women with a mother or sister who was diagnosed with breast cancer (48, 59).

Theories to explain the transient increased risk of breast cancer after childbirth vary. High levels of estrogen and progesterone and the rapid expansion of breast cells during pregnancy could promote latent initiated tumor cells. However, breast tumors diagnosed postpartum are more often at an advanced stage and are associated with lower survival compared with those diagnosed during pregnancy (67–69). This evidence has led to increased focus on the role of post-partum exposures, including lactational involution (the process that returns the mammary gland to a non-milk producing state), as contributors to a pro-tumorigenic microenvironment that may be favorable for cancer cell migration and metastasis (70). Potential adverse effects of lactational involution on the breast microenvironment must also be reconciled with demonstrated lower risks of specific tumor subtypes among parous women who breastfeed, including ER-negative or basal-like tumors that confer a worse prognosis (56, 65). A better understanding of the factors that contribute to short-term increases in breast cancer risk after pregnancy, including potential variation by age, parity, oral contraceptive use, breastfeeding, family history, or tumor subtype could provide necessary information for refining hypotheses about carcinogenesis in reproductive-age women (71). Individual studies have had insufficient statistical power or have lacked key information to evaluate these

characteristics jointly, making the Collaborative Group an ideal setting in which to advance understanding of pregnancy's role in premenopausal breast cancer development.

BMI and other anthropometrics

There is epidemiological evidence for higher BMI at premenopausal ages having an inverse association with breast cancer risk (72–75). Higher adiposity in childhood and adolescence appears to be associated with a lower risk of breast cancer at both premenopausal (73, 76–78) and postmenopausal (77–79) ages. Whether further weight gain contributes to additional reductions in premenopausal breast cancer risk is not entirely clear (80, 81). A protective effect of adiposity at premenopausal ages is in contrast with the effect of adiposity at postmenopausal ages, with greater BMI after menopause associated with higher risk of breast cancer, probably through production of estrogens by aromatase in adipose tissue (82).

The reason for the protective effect of adiposity at premenopausal ages is unclear, although several hypotheses have been put forward. Fewer ovulatory cycles in heavier women, and consequent lower sex hormone levels, has been suggested as a potential explanation (83). Similarly, an effect of polycystic ovary syndrome (PCOS) has been proposed, although Nurses' Health Study II data did not support this (73). To find the reasons for the inverse associations with premenopausal adiposity, large study populations are needed to produce stable estimates and to stratify by potentially explanatory factors.

Few published studies have had sufficiently large numbers of premenopausal cases to produce age-specific estimates over a range of ages, or to explore whether risks differ by other explanatory factors or by breast cancer subtype. The few that stratified by established breast cancer risk factors such as parity have so far reported risk estimates to be similar across these factors (72, 78). The association between adiposity and premenopausal breast cancer has been reported to vary by ethnicity, with strong associations in Caucasian, but not in Asian (84) or African-American (85), women, and associations are possibly stronger for ER^+ than ER^- premenopausal breast cancer (73). It is not clear what level of BMI confers the highest breast cancer risks—one study reported a non-linear association between BMI and risk, with the highest risk around 24 kg/m^2 (72).

The Collaborative Group, with its large number of cases in the pooled dataset and data on a wide range of risk factors, will be able to clarify the contribution of premenopausal adiposity to breast cancer risk, by examining which subtypes of breast cancer are affected, analyzing associations by exposures such as menstrual factors, and by assessing the effect of changes in adiposity over time.

Physical activity

Physical activity is of particular interest in that it constitutes a potentially modifiable risk factor for breast and other cancers. For premenopausal women, the effect of physical activity on reducing breast cancer risk appears to be smaller and less certain than for postmenopausal women (86). However, very few studies (35, 87, 88) have published prospective data for premenopausal breast cancer risk in relation to physical activity, whereas others have published by age at breast cancer (89–91) or menopausal status at study entry (92–95), or have included premenopausal women in their study but did not publish effect estimates for these women separately (96, 97).

The biological mechanisms through which physical activity could exert an effect in premenopausal women is less clear than in postmenopausal women, but might be through an effect on menarche, menstrual dysfunction, cycle length, endogenous hormone levels or estrogen metabolism (98–100). A smaller effect of physical activity in premenopausal than postmenopausal women is possible because, in contrast to postmenopausal women, in whom the protective effect of physical activity on breast cancer risk is partly through its effect on reducing adiposity, adiposity in premenopausal women has a protective effect on breast cancer risk. In addition, the impact of physical activity on hormone levels might be less influential among premenopausal women given their high levels of circulating hormones.

To aid prevention, information is needed on the type, frequency and intensity of exercise required to influence breast cancer risk, as well as the ages and characteristics of women for whom it is most effective. There might be periods of life during which physical activity has a higher impact than others, such as the time-period between menarche and first birth (101). There is also emerging evidence of differential effects of activity by ethnicity, weight, parity and family history of breast cancer, but mostly based on data from postmenopausal women (35, 91, 102). It is a limitation, however, that physical activity information is collected in many different ways and is difficult to harmonize (103).

The Collaborative Group aims to address premenopausal breast cancer risk by frequency, intensity, type, and ages of exercise, within strata defined by factors such as BMI, family history of breast cancer and age at diagnosis, and to explore specific breast cancer subtypes and stages, on a much larger scale than previously. The information gained can be used to advise young women about the extent and type of exercise that can influence their breast cancer risk.

Opportunities and challenges

The Collaborative Group is an international collaboration formed to address etiological factors for breast cancer that may be particular to, or differ in, premenopausal or perimenopausal women. By harmonizing a wide range of exposure variables across 20 studies and developing quality assurance and analysis programs, our collaboration is in a position to conduct initial analyses of pregnancy, obesity and physical activity, and to leverage the research infrastructure and established collaboration model for investigations of other risk factors. Our initial aims do not require the use of biospecimens. However, biospecimens have been collected in many of the participating studies, as described in the Cancer Epidemiology Descriptive Cohort Database (available at https://cedcd.nci.nih.gov/biospecimen.aspx) and could potentially be incorporated to address future hypotheses.

Some limitations and challenges have emerged. As in many consortia, information from the participating studies in the Collaborative Group was not collected with future pooling efforts in mind and follow-up data are not collected at standardized intervals. Therefore, harmonization efforts must identify common data elements that are collected with minimal levels of measurement error. Identification of these elements can be complicated by questionnaires and codebooks that must be translated to a common language.

Another aspect of working on pooling cohorts that requires planning and forethought is the potential for overlap of

participants between studies, for example, in Scandinavian countries with multiple cohorts that have wide geographic catchment areas. Although the existence of national identifiers makes it theoretically possible to identify women who may contribute information to more than one study in a country, the logistics for obtaining approval and merging datasets can be prohibitive. Therefore, we have worked with study investigators to identify the individual cohorts within a country with the most relevant information for specific Collaborative Group aims, and to develop strategies for excluding specific geographic regions from one cohort, but not another, where overlap of cohort catchment areas is known to exist.

The value of cancer consortia to address scientific questions efficiently and create new opportunities has become increasingly recognized (104). Conducting analyses across multiple studies requires ongoing communication and transparency. Our Collaborative Group holds in-person working group meetings in conjunction with the NCI Cohort Consortium annual meeting, as well as regular telephone conferences. These meetings provide a forum to discuss additional hypotheses that can be addressed in the future to maximize the value of the created infrastructure. The Cohort Consortium provides valuable coordinating and communication services and dedicated time and space through the annual meeting; however, other research support for data preparation, ongoing infrastructure development, and dedicated time for collaboration remains a challenge faced across many large-scale projects. Our Collaborative Group and others continue to work to identify and streamline data sharing models to maximize productivity and collaborative opportunity.

Disclosure of Potential Conflicts of Interest

No potential conflicts of interest were disclosed.

Acknowledgments

We wish to acknowledge all study participants, staff, and participating cancer registries as well as Hoda Anton-Culver, Jianwen Cai, Jessica Clague, Christina Clarke, Dennis Deapen, Niclas Håkansson, Allison Iwan, Diane Kampa, James Lacey, Eunjung Lee, Siew-Hong Low, David Nelson, Susan Neuhausen, Katie O'Brien, Hannah Park, Jerry Reid, Peggy Reynolds, Sophia Wang, Renwei Wang, Mark Weaver, Jiawei Xu, Jeffrey Yu, and Argyrios Ziogas.

Grant Support

Support for this research comes, in part, from the Avon Foundation (02-2014-080); Breast Cancer Now; The Institute of Cancer Research, London; the United States National Institutes of Health National Institute of Environmental Health Sciences (Z01 ES044005, P30 ES000260) and National Cancer Institute (UM1 CA176726, UM1 CA186107, UM1 CA182876, UM1 CA182934, UM1 CA164974, R01 CA058420, R01 CA092447, CA077398, CA144034); the United States National Center for Advancing Translational Sciences (KL2-TR001109), the National Program of Cancer Registries of the Centers for Disease Control and Prevention, and the Department of Energy; the Swedish Research Council and Swedish Cancer Foundation; the Japanese Ministry of Health, Labor and Welfare; the Hellenic Health Foundation; Karolinska Institutet Distinguished Professor Award Dnr: 2368/10-221; Cancer Council Victoria and the Australia National Health and Medical Research Council (209057, 396414, 504711); the State of Maryland, the Maryland Cigarette Restitution Fund; and the United Kingdom National Health Service funding to the Royal Marsden/ICR NIHR Biomedical Research Centre. The coordination of the European Prospective Investigation in Cancer is financially supported by the European Commission (DG-SANCO) and the International Agency for Research on Cancer. The national cohorts are supported by Danish Cancer Society (Denmark); Ligue Contre le Cancer, Institut Gustave Roussy, Mutuelle Générale de l'Education Nationale, Institut National de la Santé et de la Recherche Médicale (INSERM; France); German Cancer Aid, German Cancer Research Center (DKFZ), Federal Ministry of Education and Research (BMBF), Deutsche Krebshilfe, Deutsches Krebsforschungszentrum and Federal Ministry of Education and Research (Germany); the Hellenic Health Foundation (Greece); Associazione Italiana per la Ricerca sul Cancro-AIRC-Italy and National Research Council (Italy); Dutch Ministry of Public Health, Welfare and Sports (VWS), Netherlands Cancer Registry (NKR), LK Research Funds, Dutch Prevention Funds, Dutch ZON (Zorg Onderzoek Nederland), World Cancer Research Fund (WCRF), Statistics Netherlands (The Netherlands); ERC-2009-AdG 232997 and Nordforsk, Nordic Centre of Excellence program on Food, Nutrition and Health (Norway); Health Research Fund (FIS), PI13/00061 to Granada, PI13/01162 to EPIC-Murcia, Regional Governments of Andalucía, Asturias, Basque Country, Murcia and Navarra, ISCIII RETIC (RD06/0020; Spain); Swedish Cancer Society, Swedish Research Council and County Councils of Skåne and Västerbotten (Sweden); Cancer Research UK (14136 to EPIC-Norfolk; C570/A16491 and C8221/A19170 to EPIC-Oxford), Medical Research Council (1000143 to EPIC-Norfolk, MR/M012190/1 to EPIC-Oxford; United Kingdom).

The costs of publication of this article were defrayed in part by the payment of page charges. This article must therefore be hereby marked *advertisement* in accordance with 18 U.S.C. Section 1734 solely to indicate this fact.

Received March 16, 2017; revised April 28, 2017; accepted May 23, 2017; published OnlineFirst June 9, 2017.

References

1. Ferlay J, Soerjomataram I, Ervik M, Dikshit R, Eser S, Mathers C, et al. GLOBOCAN 2012 v1.0, Cancer Incidence and Mortality Worldwide: IARC CancerBase No. 11 [Internet]; 2013. Available from:http://globocan.iarc.fr.
2. Rosenberg SM, Partridge AH. Management of breast cancer in very young women. Breast 2015;24:S154–8.
3. Lewis DR, Seibel NL, Smith AW, Stedman MR. Adolescent and young adult cancer survival. J Natl Cancer Inst Monogr 2014;2014:228–35.
4. Johnson RH, Chien FL, Bleyer A. Incidence of breast cancer with distant involvement among women in the United States, 1976 to 2009. JAMA 2013;309:800–5.
5. Schedin P. Pregnancy-associated breast cancer and metastasis. Nat Rev Cancer 2006;6:281–91.
6. White AJ, Nichols HB, Bradshaw PT, Sandler DP. Overall and central adiposity and breast cancer risk in the Sister Study. Cancer 2015; 121:3700–8.
7. Johnson KC, Miller AB, Collishaw NE, Palmer JR, Hammond SK, Salmon AG, et al. Active smoking and secondhand smoke increase breast cancer risk: the report of the Canadian Expert Panel on Tobacco Smoke and Breast Cancer Risk (2009). Tob Control 2011;20:e2.
8. Perou CM, Sorlie T, Eisen MB, van de Rijn M, Jeffrey SS, Rees CA, et al. Molecular portraits of human breast tumours. Nature 2000;406: 747–52.
9. Clarke CA, Keegan TH, Yang J, Press DJ, Kurian AW, Patel AH, et al. Age-specific incidence of breast cancer subtypes: understanding the black-white crossover. J Natl Cancer Inst 2012;104: 1094–101.
10. Boggs DA, Rosenberg L, Adams-Campbell LL, Palmer JR. Prospective approach to breast cancer risk prediction in African American women: the black women's health study model. J Clin Oncol 2015; 33:1038–44.
11. Swerdlow A, Jones M, Schoemaker M, Hemming J, Thomas D, Williamson J, et al. The Breakthrough Generations Study: design of a long-term UK cohort study to investigate breast cancer aetiology. Br J Cancer 2011;105:911–7.
12. Rohan TE, Soskolne CL, Carroll KK, Kreiger N. The Canadian Study of Diet, Lifestyle, and Health: design and characteristics of a new cohort study of cancer risk. Cancer Detect Prev 2007;31:12–7.
13. Gallicchio L, Visvanathan K, Burke A, Hoffman SC, Helzlsouer KJ. Nonsteroidal anti-inflammatory drugs and the risk of developing breast

cancer in a population-based prospective cohort study in Washington County, MD. Int J Cancer 2007;121:211–15.
14. Riboli E, Hunt K, Slimani N, Ferrari P, Norat T, Fahey M, et al. European Prospective Investigation into Cancer and Nutrition (EPIC): study populations and data collection. Public Health Nutr 2002;5:1113–24.
15. Clavel-Chapelon F, E3N Study Group. Cohort profile: the French E3N cohort study. Int J Epidemiol 2014:dyu184.
16. Giles G, English D, Riboli E, Lambert R. The Melbourne Collaborative Cohort Study. Nutrition and lifestyle: opportunities for cancer prevention European Conference on Nutrition and Cancer held in Lyon, France in 21–24 June, 2003: International Agency for Research on Cancer (IARC). p. 69–70. Available from: http://www.iarc.fr/en/publications/pdfs-online/prev/sp156/sp156-ch1.pdf.
17. Hennekens C, Speizer F, Rosner B, Bain C, Belanger C, Peto R. Use of permanent hair dyes and cancer among registered nurses. Lancet 1979; 313:1390–3.
18. Colditz GA, Hankinson SE. The Nurses' Health Study: lifestyle and health among women. Nat Rev Cancer 2005;5:388–96.
19. Toniolo PG, Levitz M, Zeleniuch-Jacquotte A, Banerjee S, Koenig KL, Shore RE, et al. A prospective study of endogenous estrogens and breast cancer in postmenopausal women. J Natl Cancer Inst 1995; 87:190–7.
20. Zeleniuch-Jacquotte A, Afanasyeva Y, Kaaks R, Rinaldi S, Scarmo S, Liu M, et al. Premenopausal serum androgens and breast cancer risk: a nested case-control study. Breast Cancer Res 2012;14:R32.
21. Beebe G, Ishida M, Jablon S. Description of study mortality in the medical subsample October 1950–June 1951. J Hiroshima Med Assoc 1962;15: 1397–422.
22. Hankin JH, Stram DO, Arakawa K, Park S, Low S-H, Lee H-P, et al. Singapore Chinese Health Study: development, validation, and calibration of the quantitative food frequency questionnaire. Nutr Cancer 2001;39:187–95.
23. Signorello LB, Hargreaves MK, Steinwandel MD, Zheng W, Cai Q, Schlundt DG, et al. Southern community cohort study: establishing a cohort to investigate health disparities. J Natl Med Assoc 2005;97: 972.
24. Wolk A, Bergström R, Hunter D, Willett W, Ljung H, Holmberg L, et al. A prospective study of association of monounsaturated fat and other types of fat with risk of breast cancer. Arch Intern Med 1998;158:41–5.
25. Doody MM, Freedman DM, Alexander BH, Hauptmann M, Miller JS, Rao RS, et al. Breast cancer incidence in U.S. radiologic technologists. Cancer 2006;106:2707–15.
26. Krokstad S, Langhammer A, Hveem K, Holmen TL, Midthjell K, Stene TR, et al. Cohort Profile: the HUNT Study, Norway. Int J Epidemiol 2013;42:968–77.
27. Roswall N, Sandin S, Adami HO, Weiderpass E. Cohort Profile: the Swedish women's lifestyle and health cohort. Int J Epidemiol. 2015 Jun 10. [Epub ahead of print].
28. Bernstein L, Allen M, Anton-Culver H, Deapen D, Horn-Ross PL, Peel D, et al. High breast cancer incidence rates among California teachers: results from the California Teachers Study (United States). Cancer Causes Control 2002;13:625–35.
29. Toniolo PG, Pasternack BS, Shore RE, Sonnenschein E, Koenig KL, Rosenberg C, et al. Endogenous hormones and breast cancer: a prospective cohort study. Breast Cancer Res Treat 1991;18:S23–S6.
30. Gertig DM, Fletcher AS, English DR, MacInnis RJ, Hopper JL, Giles GG. Hormone therapy and breast cancer: what factors modify the association? Menopause 2006;13:178–84.
31. McGregor DH, Land C, Choi K, Tokuoka S, Liu PI, Wakabayashi T, et al. Breast cancer incidence among atomic bomb survivors, Hiroshima and Nagasaki, 1950–69. J Natl Cancer Inst 1977;59:799–811.
32. Gago-Dominguez M, Yuan J, Sun C, Lee H, Yu M. Opposing effects of dietary n-3 and n-6 fatty acids on mammary carcinogenesis: The Singapore Chinese Health Study. Br J Cancer 2003;89:1686–92.
33. Eiliv L, Merethe K, Tonje B, Anette H, Kjersti B, Elise E, et al. External validity in a population-based national prospective study–the Norwegian Women and Cancer Study (NOWAC). Cancer Causes Control 2003; 14:1001–8.
34. Willett WC, Stampfer MJ, Colditz GA, Rosner BA, Hennekens CH, Speizer FE. Dietary fat and the risk of breast cancer. N Engl J Med 1987;316:22–8.
35. Cho E, Spiegelman D, Hunter DJ, Chen WY, Stampfer MJ, Colditz GA, et al. Premenopausal fat intake and risk of breast cancer. J Natl Cancer Inst 2003;95:1079–85.
36. Harville DA.Maximum likelihood approaches to variance component estimation and to related problems. J Am Stat Assoc 1977; 72:320–38.
37. Laird NM, Ware JH. Random effects models for longitudinal data. Biometrics 1982;38:963–74.
38. Smith-Warner SA, Spiegelman D, Ritz J, Albanes D, Beeson WL, Bernstein L, et al. Methods for pooling results of epidemiologic studies: the pooling project of prospective studies of diet and cancer. Am J Epidemiol 2006;163:1053–64.
39. Cochran WG.The combination of estimates from different experiments. Biometrics 1954;10:101–29.
40. DerSimonian R, Laird N. Meta-analysis in clinical trials. Control Clin Trials 1986;7:177–88.
41. Higgins J, Thompson S, Deeks J, Altman D. Statistical heterogeneity in systematic reviews of clinical trials: a critical appraisal of guidelines and practice. J Health Serv Res Policy 2002;7:51–61.
42. Cox DR.Regression models and life-tables. J Royal Stat Soc 1972;34: 187–220.
43. Xue X, Kim MY, Gaudet MM, Park Y, Heo M, Hollenbeck AR, et al. A comparison of the polytomous logistic regression and joint cox proportional hazards models for evaluating multiple disease subtypes in prospective cohort studies. Cancer Epidemiol Biomarkers Prev 2013;22:275–85.
44. Lunn M, McNeil D. Applying Cox regression to competing risks. Biometrics 1995;51:524–32.
45. Lambe M, Hsieh C, Trichopoulos D, Ekbom A, Pavia M, Adami HO. Transient increase in the risk of breast cancer after giving birth. N Engl J Med 1994;331:5–9.
46. Albrektsen G, Heuch I, Kvale G. The short-term and long-term effect of a pregnancy on breast cancer risk: a prospective study of 802,457 parous Norwegian women. Br J Cancer 1995;72:480–4.
47. Leon DA, Carpenter LM, Broeders MJ, Gunnarskog J, Murphy MF. Breast cancer in Swedish women before age 50: evidence of a dual effect of completed pregnancy. Cancer Causes Control 1995;6:283–91.
48. Wohlfahrt J, Olsen JH, Melby M. Breast cancer risk after childbirth in young women with family history (Denmark). Cancer Causes Control 2002;13:169–74.
49. Kauppila A, Kyyronen P, Lehtinen M, Pukkala E. Dual effect of short interval between first and second birth on ductal breast cancer risk in Finland. Cancer Causes Control 2012;23:187–93.
50. Williams EM, Jones L, Vessey MP, McPherson K. Short term increase in risk of breast cancer associated with full term pregnancy. BMJ 1990; 300:578–9.
51. Bruzzi P, Negri E, La Vecchia C, Decarli A, Palli D, Parazzini F, et al. Short term increase in risk of breast cancer after full term pregnancy. BMJ 1988;297:1096–8.
52. Hsieh C, Pavia M, Lambe M, Lan SJ, Colditz GA, Ekbom A, et al. Dual effect of parity on breast cancer risk. Eur J Cancer 1994;30A:969–73.
53. Chie WC, Hsieh C, Newcomb PA, Longnecker MP, Mittendorf R, Greenberg ER, et al. Age at any full-term pregnancy and breast cancer risk. Am J Epidemiol 2000;151:715–22.
54. Cummings P, Weiss NS, McKnight B, Stanford JL. Estimating the risk of breast cancer in relation to the interval since last term pregnancy. Epidemiology 1997;8:488–94.
55. Palmer JR, Viscidi E, Troester MA, Hong CC, Schedin P, Bethea TN, et al. Parity, lactation, and breast cancer subtypes in African American women: results from the AMBER Consortium. J Natl Cancer Inst 2014;106:pii: dju237.
56. Palmer JR, Boggs DA, Wise LA, Ambrosone CB, Adams-Campbell LL, Rosenberg L. Parity and lactation in relation to estrogen receptor negative breast cancer in African American women. Cancer Epidemiol Biomarkers Prev 2011;20:1883–91.
57. Borges VF, Schedin PJ. Pregnancy-associated breast cancer: an entity needing refinement of the definition. Cancer 2012;118:3226–8.
58. Kahlenborn C, Modugno F, Potter DM, Severs WB. Oral contraceptive use as a risk factor for premenopausal breast cancer: a meta-analysis. Mayo Clin Proc 2006;81:1290–302.

59. Albrektsen G, Heuch I, Thoresen S, Kvale G. Family history of breast cancer and short-term effects of childbirths on breast cancer risk. Int J Cancer 2006;119:1468–74.
60. Cruz GI, Martinez ME, Natarajan L, Wertheim BC, Gago-Dominguez M, Bondy M, et al. Hypothesized role of pregnancy hormones on HER2+ breast tumor development. Breast Cancer Res Treat 2013; 137:237–46.
61. Albrektsen G, Heuch I, Thoresen SO. Histological type and grade of breast cancer tumors by parity, age at birth, and time since birth: a register-based study in Norway. BMC Cancer 2010;10:226.
62. Albrektsen G, Heuch I, Thoresen S, Kvale G. Clinical stage of breast cancer by parity, age at birth, and time since birth: a progressive effect of pregnancy hormones? Cancer Epidemiol Biomarkers Prev 2006; 15:65–9.
63. Innes KE, Byers TE. First pregnancy characteristics and subsequent breast cancer risk among young women. Int J Cancer 2004;112:306–11.
64. Ewertz M, Duffy SW, Adami HO, Kvale G, Lund E, Meirik O, et al. Age at first birth, parity and risk of breast cancer: a meta-analysis of 8 studies from the Nordic countries. Int J Cancer 1990;46:597–603.
65. Millikan RC, Newman B, Tse CK, Moorman PG, Conway K, Dressler LG, et al. Epidemiology of basal-like breast cancer. Breast Cancer Res Treat 2008;109:123–39.
66. Work ME, John EM, Andrulis IL, Knight JA, Liao Y, Mulligan AM, et al. Reproductive risk factors and oestrogen/progesterone receptor-negative breast cancer in the Breast Cancer Family Registry. Br J Cancer 2014;110:1367–77.
67. Johansson AL, Andersson TM, Hsieh CC, Cnattingius S, Lambe M. Increased mortality in women with breast cancer detected during pregnancy and different periods postpartum. Cancer Epidemiol Biomarkers Prev 2011;20:1865–72.
68. Stensheim H, Moller B, van Dijk T, Fossa SD. Cause-specific survival for women diagnosed with cancer during pregnancy or lactation: a registry-based cohort study. J Clin Oncol 2009;27:45–51.
69. Johansson AL, Andersson TM, Hsieh CC, Jirstrom K, Dickman P, Cnattingius S, et al. Stage at diagnosis and mortality in women with pregnancy-associated breast cancer (PABC). Breast Cancer Res Treat 2013;139: 183–92.
70. Lyons TR, O'Brien J, Borges VF, Conklin MW, Keely PJ, Eliceiri KW, et al. Postpartum mammary gland involution drives progression of ductal carcinoma in situ through collagen and COX-2. Nat Med 2011;17: 1109–15.
71. Faupel-Badger JM, Arcaro KF, Balkam JJ, Eliassen AH, Hassiotou F, Lebrilla CB, et al. Postpartum remodeling, lactation, and breast cancer risk: summary of a national cancer Institute-Sponsored Workshop. J Natl Cancer Inst 2013;105:166–74.
72. van den Brandt PA, Spiegelman D, Yaun SS, Adami HO, Beeson L, Folsom AR, et al. Pooled analysis of prospective cohort studies on height, weight, and breast cancer risk. Am J Epidemiol 2000;152:514–27.
73. Michels KB, Terry KL, Willett WC. Longitudinal study on the role of body size in premenopausal breast cancer. Arch Intern Med 2006; 166:2395–402.
74. Tamimi RM, Colditz GA, Hazra A, Baer HJ, Hankinson SE, Rosner B, et al. Traditional breast cancer risk factors in relation to molecular subtypes of breast cancer. Breast Cancer Res Treat 2012; 131:159–67.
75. Guo Y, Warren Andersen S, Shu XO, Michailidou K, Bolla MK, Wang Q, et al. Genetically predicted body mass index and breast cancer risk: mendelian randomization analyses of data from 145,000 Women of European Descent. PLoS Med 2016;13:e1002105.
76. Huang Z, Hankinson SE, Colditz GA, Stampfer MJ, Hunter DJ, Manson JE, et al. Dual effects of weight and weight gain on breast cancer risk. JAMA 1997;278:1407–11.
77. Berkey CS, Frazier AL, Gardner JD, Colditz GA. Adolescence and breast carcinoma risk. Cancer 1999;85:2400–9.
78. Baer HJ, Tworoger SS, Hankinson SE, Willett WC. Body fatness at young ages and risk of breast cancer throughout life. Am J Epidemiol 2010;171:1183–94.
79. Bardia A, Vachon CM, Olson JE, Vierkant RA, Wang AH, Hartmann LC. Relative weight at age 12 and risk of postmenopausal breast cancer. Cancer Epidemiol Biomarkers Prev 2008;17:374–8.
80. Coates RJ, Uhler RJ, Hall HI, Potischman N, Brinton LA, Ballard-Barbash R, et al. Risk of breast cancer in young women in relation to body size and weight gain in adolescence and early adulthood. Br J Cancer 1999; 81:167–74.
81. Rosner B, Eliassen AH, Toriola AT, Chen WY, Hankinson SE, Willett WC, et al. Weight and weight changes in early adulthood and later breast cancer risk. Int J Cancer 2017;140:2003–2014.
82. Key TJ, Appleby PN, Reeves GK, Roddam A, Dorgan JF, Longcope C, et al. Body mass index, serum sex hormones, and breast cancer risk in postmenopausal women. J Natl Cancer Inst 2003; 95:1218–26.
83. Key TJ, Pike MC. The role of oestrogens and progestagens in the epidemiology and prevention of breast cancer. Eur J Cancer Clin Oncol 1988;24:29–43.
84. Amadou A, Ferrari P, Muwonge R, Moskal A, Biessy C, Romieu I, et al. Overweight, obesity and risk of premenopausal breast cancer according to ethnicity: a systematic review and doseresponse metaanalysis. Obesity Rev 2013;14:665–78.
85. Robinson WR, Tse CK, Olshan AF, Troester MA. Body size across the life course and risk of premenopausal and postmenopausal breast cancer in Black women, the Carolina Breast Cancer Study, 1993–2001. Cancer Causes Control 2014;25:1101–17.
86. Friedenreich CM.Physical activity and breast cancer: review of the epidemiologic evidence and biologic mechanisms. Recent Results Cancer Res 2011;188:125–39.
87. Maruti SS, Willett WC, Feskanich D, Rosner B, Colditz GA. A prospective study of age-specific physical activity and premenopausal breast cancer. J Natl Cancer Inst 2008;100:728–37.
88. Rosenberg L, Palmer JR, Bethea TN, Ban Y, Kipping-Ruane K, Adams-Campbell LL. A prospective study of physical activity and breast cancer incidence in African-American women. Cancer Epidemiol Biomarkers Prev 2014;23:2522–31.
89. Thune I, Brenn T, Lund E, Gaard M. Physical activity and the risk of breast cancer. N Engl J Med 1997;336:1269–75.
90. Sesso HD, Paffenbarger RS Jr, Lee IM. Physical activity and breast cancer risk in the College Alumni Health Study (United States). Cancer Causes Control 1998;9:433–9.
91. Dallal CM, Sullivan-Halley J, Ross RK, Wang Y, Deapen D, Horn-Ross PL, et al. Long-term recreational physical activity and risk of invasive and in situ breast cancer: the California teachers study. Arch Intern Med 2007;167:408–15.
92. Lahmann PH, Friedenreich C, Schuit AJ, Salvini S, Allen NE, Key TJ, et al. Physical activity and breast cancer risk: the European Prospective Investigation into Cancer and Nutrition. Cancer Epidemiol Biomarkers Prev 2007;16:36–42.
93. Howard RA, Leitzmann MF, Linet MS, Freedman DM. Physical activity and breast cancer risk among pre- and postmenopausal women in the U.S. Radiologic Technologists cohort. Cancer Causes Control 2009; 20:323–33.
94. Margolis KL, Mucci L, Braaten T, Kumle M, Trolle LY, Adami HO, et al. Physical activity in different periods of life and the risk of breast cancer: the Norwegian-Swedish Women's Lifestyle and Health cohort study. Cancer Epidemiol Biomarkers Prev 2005;14:27–32.
95. Silvera SA, Jain M, Howe GR, Miller AB, Rohan TE. Energy balance and breast cancer risk: a prospective cohort study. Breast Cancer Res Treat 2006;97:97–106.
96. Rockhill B, Willett WC, Hunter DJ, Manson JE, Hankinson SE, Colditz GA. A prospective study of recreational physical activity and breast cancer risk. Arch Intern Med 1999;159:2290–6.
97. Tehard B, Friedenreich CM, Oppert JM, Clavel-Chapelon F. Effect of physical activity on women at increased risk of breast cancer: results from the E3N cohort study. Cancer Epidemiol Biomarkers Prev 2006;15:57–64.
98. Gaudet MM, Press MF, Haile RW, Lynch CF, Glaser SL, Schildkraut J, et al. Risk factors by molecular subtypes of breast cancer across a population-based study of women 56 years or younger. Breast Cancer Res Treat 2011;130:587–97.
99. Smith AJ, Phipps WR, Thomas W, Schmitz KH, Kurzer MS. The effects of aerobic exercise on estrogen metabolism in healthy premenopausal women. Cancer Epidemiol Biomarkers Prev 2013;22:756–64.

100. Campbell KL, McTiernan A. Exercise and biomarkers for cancer prevention studies. J Nutr 2007;137:161S–9S.
101. Liu Y, Tobias DK, Sturgeon KM, Rosner B, Malik V, Cespedes E, et al. Physical activity from menarche to first pregnancy and risk of breast cancer. Int J Cancer 2016;139:1223–30.
102. Lynch BM, Neilson HK, Friedenreich CM. Physical activity and breast cancer prevention. Recent Results Cancer Res 2011;186:13–42.
103. Moore SC, Lee IM, Weiderpass E, Campbell PT, Sampson JN, Kitahara CM, et al. Association of leisure-time physical activity with risk of 26 types of cancer in 1.44 million adults. JAMA Intern Med 2016; 176:816–25.
104. Burgio MR, Ioannidis JP, Kaminski BM, Derycke E, Rogers S, Khoury MJ, et al. Collaborative cancer epidemiology in the 21st century: the model of cancer consortia. Cancer Epidemiol Biomarkers Prev 2013;22:2148–60.
105. Lund E, Dumeaux V, Braaten T, Hjartaker A, Engeset D, Skeie G, et al. Cohort profile: the Norwegian Women and Cancer Study–NOWAC–Kvinner og kreft. Int J Epidemiol 2008;37:36–41.