Examining the Etiology of Early-Onset Breast Cancer in the Canadian Partnership for Tomorrow's Health (CanPath)

Joy Pader

Alberta Health Services

Robert B. Basmadjian

University of Calgary

Dylan E. O'Sullivan

University of Calgary

Nicole E. Mealey

University of Calgary

Yibing Ruan

Alberta Health Services

Christine Friedenreich

Alberta Health Services

Rachel Murphy

University of British Columbia

Edwin Wang

University of Calgary

May Lynn Quan

Alberta Health Services

Darren R. Brenner (✉ Darren.Brenner@ucalgary.ca)

University of Calgary <https://orcid.org/0000-0002-3027-290X>

Research Article

Keywords: early-onset breast cancer, Canadian Partnership for Tomorrow's Health (CanPath), etiology

Posted Date: March 19th, 2021

DOI: <https://doi.org/10.21203/rs.3.rs-312244/v1>

License: © ⓘ This work is licensed under a Creative Commons Attribution 4.0 International License.

[Read Full License](#)

Abstract

Purpose: Breast cancer incidence among younger women (under age 50) has increased over the past 25 years, yet little is known about the etiology among this age group. The objective of this study was to investigate relationships between modifiable and non-modifiable risk factors and early-onset breast cancer among three prospective Canadian cohorts.

Methods: A matched case-control study was conducted using data from Alberta's Tomorrow Project, BC Generations Project, and the Ontario Health Study. Participants diagnosed with breast cancer before age 50 were identified through provincial registries and matched to three control participants of similar age and follow-up. Conditional logistic regression was used to examine the association between factors and risk of early-onset breast cancer.

Results: In total, 609 cases and 1,827 controls were included. A body mass index $\geq 30 \text{ kg/m}^2$ was associated with a lower risk of early-onset breast cancer (OR=0.65; 95% CI: 0.47-0.90), while a waist circumference ≥ 88 cm was associated with an increased risk (OR=1.40; 95% CI: 1.06-1.84). A reduced risk was found for women with ≥ 2 pregnancies (OR=0.80; 95% CI: 0.64-1.00) and a first-degree family history of breast cancer was associated with an increased risk (OR=2.06; 95% CI: 1.54-2.75).

Conclusions: In this study, measures of adiposity, pregnancy history, and familial history of breast cancer are important risk factors for early-onset breast cancer. Evidence was insufficient to conclude if smoking, alcohol intake, fruit and vegetable consumption, and physical activity are meaningful risk factors. The results of this study could inform targeted primary and secondary prevention for early-onset breast cancer.

Introduction

Breast cancer is the most commonly diagnosed cancer globally among women with over two million new cases in 2018.(1) Canada is among the top 25 countries worldwide with the highest rates of breast cancer.(1) A projected 27,400 breast cancer cases are expected to be diagnosed in Canada in 2020.(2) Breast cancer incidence among younger women (under age 50) has increased in Canada since 2000.(3) The Canadian Cancer Society (CCS) projected that in 2017, roughly 17% of new breast cancer cases will occur before age 50.(4) Breast cancer diagnosed among younger women often occur outside age-restricted or high-risk screening programs, resulting in advanced stage at diagnosis and poorer survival rates.(5, 6) The diagnosis of early-onset breast cancer also presents several unique challenges, including fertility concerns and reduced quality of life due to treatment-induced premature menopause.(5, 6) While mutations in the *BRCA1* and *BRCA2* genes increase the risk of developing breast cancer, only a small number of breast cancers (~ 5–10%) diagnosed among young women are attributable to these mutations,(7, 8) suggesting other genetic, environmental or lifestyle factors may contribute to the development of early-onset breast cancer.

Several risk factors for breast cancer have been identified, with the majority of these risk factors applying to both pre and post-menopausal breast cancer.(9) While these risk factors for old and younger breast cancer cases are similar, there is evidence of early-onset breast cancer being biologically and etiologically distinct from breast cancer diagnosed among woman of older ages and further research is required.(10). Specifically, the relationship between anthropometric factors and early-onset breast cancer risk requires clarification, with previously reported inverse associations and a conflicted evidence base.(11) For example, current evidence suggests obesity is a risk factor for postmenopausal breast cancer(12, 13) while being protective for premenopausal breast cancer.(14) Furthermore, the mechanisms explaining the underlying pathways in these associations remain unclear. However, present studies on early-onset breast cancer and associated risk factors have been limited or conflicting.(10, 15, 16) Studies that have attempted to resolve previous conflicted evidence but have been impacted by smaller comparative sample sizes among younger women and varied definitions of “young women” across studies.(10, 15, 16)

The objective of our study was to examine the impact of lifestyle and reproductive factors, as well as family and medical history on the risk of developing early-onset breast cancer (under the age of 50) using data from The Canadian Partnership for Tomorrow’s Health (CanPath).(17) The CanPath cohort is a national prospective cohort study that was developed to explore the relationships between various lifestyle, genetic, and environmental factors and outcomes of disease. The national study population includes over 330,000 Canadians who were recruited between the ages of 30 and 74 years from six regional cohorts. We combined three of the six cohorts involved in CanPath: Alberta’s Tomorrow Project (ATP); British Columbia Generations Project (BCGP); and the Ontario Health Study (OHS) to investigate various risk factors for early-onset breast cancer.

Materials And Methods

Study Population

In this study we pooled baseline data from 2,437 participants from the ATP, OHS and BCGP regional cohorts. ATP was initiated in 2000 and enrolled approximately 55,000 participants between the ages of 35–69 at three different time points – 2001 (phase 1), 2008 (phase 2), and 2011 (phase 3). Data from phases 2 and 3 from ATP were used in this analysis since the questionnaires from those phases were harmonized with the other cohorts in CanPath. OHS has recruited over 225,000 Ontario residents in the age range of 35–74 starting in 2009, while BCGP has enrolled approximately 30,000 British Columbia residents between the ages of 30 and 74 starting in 2009.

Study Design

A nested matched case-control study design was used for this study.

Cases and controls were selected from the same regional cohort in a 1:3 case-control ratio. Cases were defined as female participants who had an incident, primary breast cancer diagnosis under 50 years of age during the follow-up period. The incident cases were identified through data linkages done with the BC Cancer Registry, Alberta Cancer Registry and Cancer Care Ontario, respectively, using

participants' Personal Health Identification Numbers. Using an incidence density sampling approach, controls were sampled without replacement and matched to cases on follow-up (± 6 months) and age (± 1 year). The follow-up period for each cohort was defined from cohort initiation to the most recent linkage with their respective provincial cancer registry, which was 2008–2018 for Alberta, 2009–2017 for British Columbia, and 2009–2017 for Ontario.

Data Collection – Questionnaires

Data on demographic, lifestyle, reproductive, and family and medical history factors under investigation as exposures in this study were collected through self-report questionnaires completed by participants within each cohort at baseline upon cohort entry, prior to diagnoses. The exposures were analyzed as categorical variables. Thresholds for exposure categories were defined prior to analysis and informed by previous literature. Where necessary, categories were combined to ensure sufficient case counts across categories.

The demographic exposures included marital status (married/living with a partner, divorced/separated/widowed, single/never married); education (high school or less, some postsecondary, Bachelor's degree or higher); and total household income (<\$50,000, \$50,000 - \$100,000, >\$100,000).

The lifestyle exposures included smoking status (never smoked at least 100 cigarettes, past smoker of at least 100 cigarettes, current smoker); alcohol consumption in the previous 12 months (no, <once per week, $\geq$ once per week); alcohol binge drinking frequency in the previous 12 months (no, <once per week, $\geq$ once per week); physical activity (< 500 MET min/week, $\geq$ 500 MET min/week, MET = metabolic equivalency); fruit and vegetable consumption per day (0–2 servings of fruits or vegetables, 3–4 servings of fruits or vegetables, > 4 servings of fruits or vegetables, > 4 servings of fruits and vegetables); body mass index (BMI) (underweight or normal [$< 25 \text{ kg/m}^2$], overweight [$25\text{--}29.99 \text{ kg/m}^2$], obese [$\geq 30 \text{ kg/m}^2$]); waist circumference (normal/healthy [$< 88 \text{ cm}$], above normal/at risk [$\geq 88 \text{ cm}$]); and waist-to-hip ratio (normal/healthy [< 0.83], above normal/at risk [≥ 0.83]).

The reproductive exposures included: age at menarche (≤ 10 years, 11–13 years, ≥ 14 years); contraceptive use and duration (never, < 10 years, 10–19 years, ≥ 20 years); total number of pregnancies (0, 1, ≥ 2); age of first pregnancy (< 20 years, 20–29 years, ≥ 30 years); and use of fertility treatments (no, yes).

The family and medical history exposures included participation in mammography screening (no, yes); family history of breast cancer and any cancer (no, yes); history of any diabetes (no, yes), including type 1, type 2, and gestational diabetes; and history of other chronic conditions (no, yes), including irritable bowel disease/syndrome, Chron's disease, arthritis, high blood pressure, psoriasis, ulcerative colitis, liver cirrhosis, and systemic lupus erythematosus.

Statistical Analysis

Participants from each provincial cohort were pooled to generate one combined cohort. All demographic, lifestyle, reproductive, and family and medical history exposures were described using means and

standard deviations (SD) for numeric variables and frequency tables with proportions for categorical variables for the combined cohort. Multivariable conditional logistic regression models were used to evaluate the association between each exposure and the development of early-onset breast cancer. For each exposure, a multivariable model was built using the following procedure: 1) The exposure of interest and all other exposures (covariates) were included as independent variables in the conditional logistic regression model with early-onset breast cancer as the dependent variable; 2) While forcing the exposure of interest to remain in the model, backwards deletion with a liberal inclusion criterion of P-value < 0.20 was used to select covariates remaining in the model based on their association with early-onset breast cancer; 3) The exposure of interest was modeled as the dependent variable and all covariates were included as independent variables and backwards deletion with a inclusion criterion of P-value < 0.20 was used to identify covariates associated with the exposure of interest. 4) Any covariate removed from step 2 that was associated with the exposure of interest in step 3 was included in the final multivariable to ensure that any covariate that potentially confounded the association between the exposure of interest and early-onset breast cancer was adjusted for. Measures of association were quantified by the odds ratio (OR) and corresponding 95% confidence interval (95% CI). If two variables were highly correlated, we used the variable most relevant to etiology to reduce the potential of multicollinearity. In lieu of missing data, we performed complete case analyses.

Results

A total of 610 cases and 1,827 controls were eligible for analysis; however, 1 case was not matched to 3 controls and was excluded as a result. Therefore, the pooled analysis included 609 cases and 1,827 controls for a total sample size of 2,436 participants across these three regional Canadian cohorts of which 14.8% (n = 360) were from Alberta, 5.7% (n = 140) from British Columbia, and 79.5% (n = 1936) from Ontario (Table 1). The mean age of all participants at baseline was 43.1 (SD = 5.1) years. The study population had a high annual average annual income (41.2% $\geq$ \$100,000), were well-educated (44.5% had post-secondary education), and married/living with a partner (71.7%). Most of the study participants had at least one pregnancy (14.7% had one pregnancy, 63.6% had two or more), and the mean age of first pregnancy among these women was 26.1 years. In addition, these participants also reported using hormonal contraceptives (86.2%) and the mean duration of hormonal contraceptive use was 10.7 years. Only 11% of women in this study had a positive family history of breast cancer. Majority of women were absent of any diabetes (93.3%) and other chronic conditions (64.6%). Over one-third of the population was either classified as overweight (21.1%) or obese (17%) at baseline. With respect to lifestyle factors, most of study population consumed alcohol (89.6%), 17.8% were current smokers, 22.5% were past smokers, and 56.4% had never smoked. The frequency of missing data in each characteristic is also presented in Table 1.

Population characteristics of early-onset breast cancer cases and controls, Alberta's w Project (2008-2018), BC Generations Project (2009-2017), Ontario Health Study (2009-2018)
n=2436

Characteristics	Categories	Cases (n =609)	Controls (n =1827)	Total (n =2436)
	Alberta	90	270	360
	British Columbia	35	105	140
	Ontario	484	1452	1936
Time of survey)	Mean age, (SD)	43.2 (5.8)	43.0 (5.8)	43.1 (5.8)
Time of diagnosis)	Mean age, (SD)	43.7 (5.1)	-	-
Marital status	Married/living with a partner	426	1321	1747
	Divorced, separated, widowed	87	242	329
	Single, never married	75	237	312
	Missing	21	27	48
Education	High school or less	108	295	403
	Some post-secondary	235	685	920
	Bachelor's or higher	257	827	1084
	Missing	<10	20	29
Household Income (\$CDN)	<\$50,000	76	345	421
	\$50,000 to less than \$100,000	177	518	695
	>\$100,000	260	743	1003
	Missing	96	221	317
Menarche (years)	Mean age, y (SD)	12.7 (1.5)	12.8 (1.5)	12.7 (1.5)
	≥14	146	464	610
	11-13	393	1160	1553
	≤10	33	90	123
	Missing	37	113	150
Use of oral contraceptives	No	66	231	297
	Yes	531	1569	2100
	Missing	12	27	39
Duration of oral contraceptives	Mean years, y (SD)	10.4 (6.8)	10.7 (7.4)	10.6 (7.3)
	Never used	66	237	303
	Less than 10 years	216	637	853
	10 to 19 years	182	526	708
	20 or more years	73	209	282
	Missing	72	218	290
Number of pregnancies	Mean number (SD)	2.1 (2.0)	2.3 (1.8)	2.2 (1.9)
	0	131	357	488
	1	101	256	357
	2 to 3	271	829	1100
	≥4	89	352	441
	Missing	17	33	50
Age at first pregnancy (years)	Mean age at first pregnancy, (SD)	26.4 (5.9)	26.1 (6.1)	26.1 (6.1)
	Never	131	357	488
	≤19	71	237	308
	20-24	103	355	458
	25-29	146	419	565

	≥30	136	408	544
	Missing	22	51	73
and fertility treatment	No	546	1635	2181
	Yes	40	152	192
	Missing	23	40	63
had a mammogram	No	134	827	961
	Yes	453	988	1441
	Missing	22	12	34
history of breast cancer	No	475	1536	2011
	Yes	99	170	269
	Missing	35	121	156
history of any cancer	No	253	901	1154
	Yes	305	836	1141
	Missing	51	90	141
smoked (100 cigarettes)	No	328	1054	1382
	Yes	245	733	978
	Missing	36	40	76
smoking status	Never smoked	328	1054	1382
	Past smoker	181	482	663
	Current smoker	63	248	311
	Missing	37	43	80
consumed alcohol	Yes	537	1684	2221
	No	37	108	145
	Missing	35	35	70
drinking frequency last 12 months	Non-drinker	61	216	277
	Less than 1/month	121	386	507
	About once a month	57	173	230
	2-3 times/month	73	272	345
	Once a week	66	225	291
	2-3 times /week	113	273	386
	4 or more a week	83	232	315
	Missing	35	50	85
drinking frequency in the months	Non-drinker	229	682	911
	1-5 times/yr	166	529	695
	6-11 times/yr	40	101	141
	About once a month	27	95	122
	2-3 times/month	23	71	94
	Once a week	17	71	88
	2-3 times /week	22	44	66
	4-5 times /week	<10	12	17
	6-7 times /week	<10	<10	<10
	Missing	78	216	294
physical activity	Mean physical activity, MET min/wk (SD)	2309.5 (2511.9)	2425 (2723.9)	2397 (2673.2)
	<500 MET min/week	137	395	532
	≥500 MET min/week	391	1228	1619
	Missing	81	204	285
total vegetable consumption per	0-2 servings of fruit and/or vegetables	222	653	875

	3-4 servings of fruit and/or vegetables	249	822	1071
	≥4 servings of either fruit or vegetables	73	243	316
	≥4 servings of fruit and vegetables	21	41	62
	Missing	44	68	112
Mass Index (BMI)	Mean BMI	26.0 (5.6)	26.9 (6.6)	26.7 (6.3)
	Underweight or Normal	235	679	914
	Overweight	121	393	514
	Obese	90	324	414
	Missing	163	431	594
Waist Circumference (cm)	Mean Waist Circumference (SD)	88.0 (13.9)	86.6 (14.4)	87 (14.3)
	Normal, Healthy (<88.0 cm)	166	593	759
	Above normal, At risk (≥88.0cm)	160	446	606
	Missing	283	788	1071
Waist-to-Hip Ratio	Mean Waist-to-Hip Ratio (SD)	0.86 (0.1)	0.85 (0.1)	0.86 (0.1)
	Normal, Healthy (<0.83)	101	375	476
	Above normal, At risk (≥0.83)	213	614	827
	Missing	295	838	1133
History of any diabetes (includes type 1 and type 2)	No	556	1718	2274
	Yes	25	78	103
	Missing	28	31	59
History of other chronic conditions	No	382	1192	1574
	Yes	214	631	845
	Missing	13	<10	17

Abbreviations: SD = standard deviation; METs = metabolic equivalent tasks (defined as the caloric need per kilogram of body weight per hour of activity divided by the caloric need per kilogram of body weight per hour)

*Includes occasional and daily smokers as current smokers.

The estimated adjusted associations between each exposure and the incidence of early-onset breast cancer in the combined cohort are presented in Table 2. The multivariable regression models used to estimate the association between each exposure and early-onset breast cancer are presented in the Supplement. There was evidence that early-onset breast cancer cases were less likely to have an average annual household income <\$50,000 vs. ≥\$100,000 (OR = 0.52; 95% CI: 0.37–0.72), to be classified as overweight vs. underweight or normal BMI (OR = 0.77; 95% CI: 0.58-1.00), obese vs. underweight or normal BMI (OR = 0.65; 95% CI: 0.47–0.90), and to have had at least two pregnancies vs. no pregnancies (OR = 0.80; 95% CI: 0.64-1.00). There was suggestive evidence that cases were less likely to be current smokers vs. never having smoked at least 100 cigarettes (OR = 0.75; 95% CI: 0.53–1.05), although this estimate did not achieve statistical significance. Early-onset breast cancer cases were more likely to have a waist circumference of ≥ 88 cm vs. <88 cm (OR = 1.40; 95% CI: 1.06–1.84), waist-to-hip ratio of ≥ 0.83 vs. <0.83 (OR = 1.39; 95% CI: 1.05–1.84), highest education of high school degree or less vs. bachelor's degree or more (OR = 1.30; 95% CI: 1.00-1.86), and a family history of breast cancer vs. no family history

(OR = 2.06; 95% CI: 1.54–2.75). There was suggestive evidence that cases were more likely to be divorced, separated, or widowed vs. married/living with a partner (OR = 1.34; 95% CI: 0.99–1.80), although this estimate did not achieve statistical significance.

Abbreviations: BMI = body mass index. CI = confidence interval. METs = metabolic equivalent tasks (defined as the caloric need per kilogram of body weight per hour of activity divided by the caloric need per kilogram of body weight per hour at rest). OR = odds ratio. *Includes occasional and daily smokers as current smokers.

Discussion

To our knowledge, this pooled analysis is the first to investigate the etiology of early-onset breast cancer in a large Canadian population. For breast cancer under 50 years, we observed risk decreases of 23%, 35%, 48% and 20% in overweight BMI, obese BMI, low income, and multiple pregnancy categories, respectively. Conversely, risk increases of 40%, 39%, 30%, and 106% were observed for larger waist circumference, larger waist-to-hip ratio, lower educational attainment, and a positive family history of breast cancer categories, respectively.

It is well established that weight gain and obesity increase breast risk in postmenopausal women.(13) Consistent positive associations have been reported between BMI, waist circumference, and waist-to-hip ratio and the risk of breast cancer among postmenopausal women(18, 19) but results are inconsistent for younger women. Our study suggests that an overweight and obese BMI are associated with lower early-onset breast cancer risk. Similar results were seen in a pooled analysis of 19 prospective studies including women under 55 years of age.(11) Our results also suggest that high waist circumference and waist-to-hip ratio are predictors of breast cancer in younger women, independent of BMI, consistent with other prospective studies adjusting for BMI in premenopausal women.(19–23) The opposing effect of BMI and waist circumference on breast cancer risk in young women seen in our study and elsewhere requires further clarification. Possible differences in these associations could be related to differential roles of overall adiposity, captured by BMI, and central adiposity, captured by waist circumference, on metabolism and their contribution to breast cancer development among younger women. Both overall and central adiposity are associated with more anovulatory cycles and lower estradiol levels in premenopausal women,(24) which would be expected to reduce breast cancer risk. However, central adiposity is also an independent predictor of both hyperinsulinemia and levels of insulin-like growth factor 1 (IGF-1), which have been previously found to be related to premenopausal breast cancer risk.(18, 25) These findings suggest that chronic inflammation and metabolic abnormalities induced by central adiposity are mechanistically important for higher breast cancer risk in young women, independent of sex hormones, which may explain the observed increase in risk for high waist circumference and waist-to-hip ratio in our study. These findings may also suggest that, in the absence of central obesity and its metabolic abnormalities, lower sex hormones are protective against breast cancer development in younger women, which could explain why young women with higher BMI are observed to be at lower breast cancer risk. Another possible explanation for the opposing effect of BMI and waist circumference

Table 2. The adjusted associations between each exposure and early-onset breast cancer in the CanPath combined cohort

Characteristics	Categories	Cases	Controls	OR (95% CI)	p-value
BMI	Underweight/normal	235	679	1.00	--
	Overweight	121	393	0.77 (0.58 - 1.00)	0.050
	Obese	90	324	0.65 (0.47 - 0.90)	0.008
Waist circumference	<88 cm	166	593	1.00	--
	≥88 cm	160	446	1.40 (1.06 - 1.84)	0.018
Waist-to-hip ratio	<0.83	101	375	1.00	--
	≥0.83	213	614	1.39 (1.05 - 1.50)	0.021
Alcohol frequency in past 12 months	Non-drinker	61	216	1.00	--
	Less than 1/week	251	831	1.03 (0.74 - 1.44)	0.859
	1/week or more	262	730	1.16 (0.82 - 1.64)	0.397
Alcohol binge drinking frequency in past 12 months	Non-drinker	229	682	1.00	--
	Less than 1/week	233	725	0.86 (0.69 - 1.09)	0.207
	1/week or more	69	204	0.96 (0.68 - 1.35)	0.795
Smoking status	Never smoker	328	1054	1.00	--
	Past smoker	181	482	1.17 (0.90 - 1.59)	0.279
	Current smoker	63	248	0.77 (0.53 - 1.06)	0.092
Marital status	Married/living with a partner	426	1321	1.00	--
	Divorced, separated, widowed	87	242	1.34 (0.99 - 1.80)	0.057
	Single, never married	75	237	1.05 (0.75 - 1.47)	0.786
Income	<\$50,000	76	345	0.52 (0.37 - 0.72)	<0.001
	≥\$50,000 <\$100,000	177	518	0.88 (0.69 - 1.12)	0.287
	≥\$100,000	260	743	1.00	--
Highest education	High school or less	108	295	1.30 (1.00 - 1.86)	0.050
	Some post-secondary	235	685	1.17 (0.93 - 1.50)	0.182
	Bachelor's degree	257	827	1.00	--
Age at menarche (years)	≥14	146	464	1.00	--
	11-13	393	1160	1.05 (0.84 - 1.31)	0.681

	≤10	33	90	1.24 (0.78 - 1.96)	0.356
Total number of pregnancies	Never	131	357	1.00	--
	1	101	256	1.06 (0.76 - 1.48)	0.784
	2 or more	360	1181	0.80 (0.64 - 1.00)	0.050
Age of first pregnancy (years)	Never	131	357	1.00	--
	≤19	71	237	0.96 (0.27 - 3.41)	0.945
	20-29	249	774	1.01 (0.29 - 3.50)	0.994
	≥30	136	408	1.06 (0.30 - 3.69)	0.932
Duration of oral contraceptives (years)	Never used	66	237	1.00	--
	<10	216	637	1.25 (0.89 - 1.74)	0.200
	10-19	182	526	1.23 (0.87 - 1.74)	0.247
	≥20	73	209	1.18 (0.74 - 1.70)	0.429
Use of fertility treatment	No	546	1635	1.00	--
	Yes	40	152	0.82 (0.57 - 1.19)	0.299
Physical activity	<500 MET min/week	137	395	1.12 (0.88 - 1.44)	0.350
	≥500 MET min/week	391	1228	1.00	--
Fruit and vegetable servings per day	≤2 servings of fruit or veg	222	653	1.00	--
	3 to 4 servings of fruit or veg	249	822	0.87 (0.66 - 1.07)	0.134
	>4 servings of fruit or veg	73	243	0.90 (0.61 - 1.18)	0.399
	>4 servings of fruit and veg	21	41	1.47 (0.84 - 2.59)	0.182
Family history of breast cancer	No	475	1536	1.00	--
	Yes	99	170	2.06 (1.54 - 2.75)	<0.0001
Family history of any cancer	No	253	901	1.00	--
	Yes	305	836	1.17 (0.93 - 1.48)	0.184
History of chronic disease	No	382	1192	1.00	--
	Yes	214	631	1.07 (0.87 - 1.31)	0.529

on breast cancer risk in young women is estrogen/progesterone (ER/PR) tumour status. High BMI has been observed to be protective in ER/PR + cancers but not in ER/PR- cancers, and higher waist circumference has been observed to increase risk of ER/PR- cancers but not ER/PR + cancers (13). The distribution of ER/PR + and ER/PR- tumours in young women in this study may account for these

opposing effects. Unfortunately, we did not have access to the tumour characteristic data to confirm whether the effect of BMI and waist circumference depended on receptor status or subtype.

A positive family history is an important risk factor for breast cancer and the strength of association increases as age of diagnosis decreases. In a pooled analysis of data from 52 epidemiological studies, including 58,209 women with breast cancer and 101,986 controls, the risk of breast cancer was two times greater in women under 50 years of age with at least one affected first-degree relative, whereas a 1.5-fold increase in risk was observed in women older than 50 years.(26) We observed an approximate two-fold risk increase in women under age 50 years with a family history of breast cancer in the Canadian population. This cumulative evidence indicates a hereditary factor for breast cancer in young women. Several germline mutations elevate breast cancer risk in young women, including the *BRCA1*, *BRCA2*, *TP53* (Li Fraumeni syndrome), and *PTEN* (Cowden's syndrome) mutations.(27) However, many young women with a strong family history of breast cancer do not carry these mutations and the genetic basis of breast cancer in young women requires further exploration.(7, 28, 29) The CCS and Canadian Task Force on Preventive Health Care (CTFPHC) recommend women with a strong family history consult with their primary care physician for personalized testing plans, which may include mammography screening at an earlier age, more frequent mammography screening, or use of an ultrasound or magnetic resonance imaging (MRI).(30, 31) The National Comprehensive Cancer Network (NCCN) guidelines recommend women at an increased risk ($\geq 20\%$ lifetime risk based on models largely dependent on family history) undergo annual MRI screening starting as early as age 25 and annual mammography screening as early as age 30.(32) This study provides additional evidence to support that young women with a positive family history of breast cancer should continue to seek counselling from their primary care provider for personalized screening strategies for early detection of breast cancer before 50 years of age. Routine mammography screening is not recommended for women aged 40–49 years in Canada as trials have shown that the number needed to screen to prevent one death from breast cancer is 1,724;(33) therefore, more efforts should be directed towards detection of high-risk young women based on familial and genetic factors.

The only reproductive risk factor that was statistically significant in this study was gravidity, in which a 20% reduction in breast cancer risk was observed for women with multiple pregnancies compared to nulliparous women. Recently, a 27% reduction in breast cancer risk was observed in a meta-analysis of 13 studies in women under the age 50 with multiple pregnancies compared to nulliparous women.(34) However, the age at first full term pregnancy influences the association between gravidity and early-onset breast cancer risk, which was not statistically significant in our models. In large prospective studies, it has been shown that women whose first full term pregnancy is in their thirties remained at an increased risk of breast cancer well into their fifties relative to nulliparous women.(35–37) Conversely, women whose first full term pregnancy is in their twenties were at an increased risk that peaks five years postpartum then declines to null 10–15 years postpartum and becomes protective around 20 years postpartum.(35, 37) Several etiologic studies explored the underlying biological processes of this association. Russo et al. found that parous postmenopausal women without breast cancer had unique gene expression patterns including differential expression of apoptosis-related genes and others related to cell cycle and cell

signaling, suggesting that pregnancy may induce a signature that provides long-term protection from developing breast cancer.(38) Another study showed that parity downregulates Wnt and Notch signaling, and suppresses progenitor cells, suggesting that this pathway could be a potential mechanism explaining the long-term protective effect of pregnancy.(39) The transient increase in the risk of breast cancer after giving birth is thought to be due to the mitogenic effect of high estrogen levels during pregnancy.(40) While our result indicates a protective effect from multiple pregnancies, no recommendations regarding risk reduction can be made given that the decision to have children is highly personal.

Most of the current evidence has found that higher socio-economic status is linked to increased risk of breast cancer, specifically among high income countries in North America and Europe.(41–46) This association is mainly seen among postmenopausal breast cancer cases.(42, 47) Our study found evidence that the odds of early-onset breast cancer were significantly lower for women with lower average annual household income. This finding is consistent with several studies that found women with higher household incomes had an increased risk of breast cancer compared to women in lower income households.(48, 49) In contrast, our study found evidence of low education (high school degree or less) associated with increased risk of early-onset breast cancer. The majority of the findings on breast cancer and education suggest that higher education is associated with increased breast cancer risk in Western or high-income countries.(46, 50, 51) Correlations of higher education and breast cancer incidence in these studies were explained by factors such as age at first pregnancy, oral contraceptive (OC) use and hormone replacement therapy use. However, results from a Dutch study found respondents with low education had an increased risk of all cancers including breast cancer among women.(52) The study suggested that factors such as alcohol consumption, obesity and poor diet may explain, in part, the association between breast cancer and low education.(52–55) Another possibility for our findings of low education and increased breast cancer risk may be related to chronic inflammation and the biomarker C-reactive protein (CRP), a protein produced by the liver in response to systemic inflammation.(56, 57) Several studies found elevated CRP levels among populations with lower education compared to higher levels of education attainment.(56, 58, 59) Relatedly, multiple systematic reviews have found chronic inflammation and CRP to be associated with increased breast cancer risk.(56, 58, 60, 61) Our study provides additional evidence on the potential life-course pathways of the impacts of socio-economic status and early-onset breast cancer.

Strengths and Limitations

Our study has numerous strengths worth highlighting.

We pooled data from a national prospective study from three regional cohorts that were each linked to provincial cancer registries.

Each cohort collected detailed health and lifestyle questionnaire data at baseline prior to a breast cancer diagnosis, ensuring that each risk factor temporally preceded the onset of disease. In addition, the prospective nature of the data ensured that the results of this study were not subject to recall or selection bias.

There are several limitations of this pooled analysis that must be acknowledged. First, data from the health and lifestyle data are from self-reported study participant questionnaires. While the questionnaires were deemed reliable and valid,⁽⁶²⁾ data from self-reported questionnaires are inherently subject to response biases such as social desirability bias. Secondly, we only had data from one time point and assumed that these data represented average exposure during young adulthood. While differential misclassification of exposure by case status is unlikely since questionnaires were answered prior to the development of breast cancer, non-differential misclassification of exposures is likely. For binary exposures this would have biased effect estimates towards the null and either towards or away from the null for exposures with three or more categories. The sample size was also insufficient to explore whether or not the association between gravidity and breast cancer risk was modified by age of first pregnancy. There were also concerns regarding missing data for several exposures, particularly BMI, waist circumference, and waist-to-hip ratio. Although missingness in exposures was unlikely related to case status and therefore would not bias our estimates, exclusion of participants with incomplete exposure resulted in losses of precision and power. Finally, we did not include race/ethnicity data for participants and could not assess whether the associations between the various exposures and early-onset breast cancer differed by racial group.

Conclusion

Our study provides further insight on sociodemographic, lifestyle, and reproductive factors associated with the risk of breast cancer under 50 years in a large Canadian prospective cohort, which may have important implications for etiology and screening. We provide evidence consistent with other epidemiologic studies that central adiposity increases breast cancer risk in women under 50 years. Lifestyle behaviours, such as a healthy diet and physical activity that reduce central adiposity may be feasible primary prevention strategies to reduce breast cancer risk, in addition to other chronic diseases in younger women. We also provide further evidence that women under 50 years with a positive family history of breast cancer should speak with their primary care provider about options for breast cancer screening, which is in accordance to screening guidelines to high-risk women set by the Canadian Cancer Society and Canadian Task Force on Preventive Health Care. Future research is still needed to explore whether or not risk factors for breast cancer under age 50 differ by molecular subtype (luminal A, luminal B, HER2 enriched, and triple-negative breast cancer). Detailed analyses are also needed to assess differences in risk factors by race and socioeconomic status, since our study only had data on income and education, but not occupation and other measures of social inequality. Etiologic research, such as exome-sequencing studies, is warranted to explore the mutational impact of lifestyle behaviours on the genome and yield discoveries regarding the underlying biological processes of breast cancer development in younger women. Such studies may provide mechanistic insights on the opposing effect of high BMI and waist circumference on breast cancer risk, as well as how pregnancy may lead to a transient increase in risk, followed by a long-term protective effect.

Declarations

Acknowledgements

We would like to thank and acknowledge the Albert Cancer Registry, Alberta's Tomorrow Project, BC Generations Project, Ontario Health Study, and Cancer Care Ontario. Alberta's Tomorrow Project is only possible because of the commitment of its research participants, its staff and its funders: Alberta Health, Alberta Cancer Foundation, Canadian Partnership Against Cancer and Health Canada, and substantial in kind funding from Alberta Health Services. Cancer registry data was obtained through linkage with Surveillance & Reporting, Cancer Research & Analytics, Cancer Care Alberta. The views expressed herein represent the views of the author(s) and not of the data sources or any of its funders. The views expressed herein represent the views of the author(s) and not of the data sources or any of its funders.

Funding Sources

Dr. Darren R. Brenner is supported by the Armstrong Investigatorship in Molecular Epidemiology from the Cumming School of Medicine at the University of Calgary and a Canadian Institutes of Health Research Grant (#397332). Dylan O'Sullivan was supported by an Ontario Graduate Scholarship and a Chen-Aronson Fellowship in Causes of Cancer.

Data Availability Statement

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

Authors' Contributions

Conception and design: D.R Brenner

Development of methodology: D.R Brenner, R.B. Basmadjian, D.E. O'Sullivan and, J. Pader

Acquisition of data (provided animals, acquired and managed patients, provided facilities, etc.): D.R Brenner, C. M. Friedenreich, R. Murphy, E. Wang, M.L.Q, and J. Pader

Analysis and interpretation of data (e.g., statistical analysis, biostatistics, computational analysis): J. Pader, R.B. Basmadjian, D.E. O'Sullivan and D.R Brenner

Writing, review, and/or revision of the manuscript: J. Pader, R.B. Basmadjian, D.E. O'Sullivan and D.R Brenner N.M., Y. Ruan., C. M. Friedenreich , R. Murphy, E. Wang, and M. L. Quan

Administrative, technical, or material support (i.e., reporting or organizing data, constructing databases): J. Pader, R.B. Basmadjian

References

1. Bray F, Ferlay J, Soerjomataram I, Siegel RL, Torre LA, Jemal A. (2018) Global cancer statistics 2018: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries. *CA: A Cancer Journal for Clinicians*. 68: 394-424.
2. Brenner DR, Weir HK, Demers AA, et al. (2020) Projected estimates of cancer in Canada in 2020. *Canadian Medical Association Journal*. 192: E199.
3. Heer E, Ruan Y, Mealey N, Quan ML, Brenner DR. (2020) The incidence of breast cancer in Canada 1971–2015: trends in screening-eligible and young-onset age groups. *Canadian Journal of Public Health*.
4. Canadian Cancer Society's Advisory Committee on Cancer Statistics. (2017) *Canadian Cancer Statistics 2017*. Toronto, ON: Canadian Cancer Society.
5. Assi HA, Khoury KE, Dbouk H, Khalil LE, Mouhieddine TH, El Saghir NS. (2013) Epidemiology and prognosis of breast cancer in young women. *J Thorac Dis*. 5 Suppl 1: S2-S8.
6. Brenner DR, Brockton NT, Kotsopoulos J, et al. (2016) Breast cancer survival among young women: a review of the role of modifiable lifestyle factors. *Cancer Causes Control*. 27: 459-72.
7. Malone KE, Daling JR, Neal C, et al. (2000) Frequency of BRCA1/BRCA2 mutations in a population-based sample of young breast carcinoma cases. *Cancer*. 88: 1393-402.
8. Peto J, Collins N, Barfoot R, et al. (1999) Prevalence of BRCA1 and BRCA2 Gene Mutations in Patients With Early-Onset Breast Cancer. *JNCI: Journal of the National Cancer Institute*. 91: 943-9.
9. Butt Z, Haider SF, Arif S, et al. (2012) Breast cancer risk factors: a comparison between pre-menopausal and post-menopausal women. *JPMA. The Journal of the Pakistan Medical Association*. 62: 120-4.
10. Chollet-Hinton L, Anders CK, Tse C-K, et al. (2016) Breast cancer biologic and etiologic heterogeneity by young age and menopausal status in the Carolina Breast Cancer Study: a case-control study. *Breast Cancer Research*. 18: 79.
11. The Premenopausal Breast Cancer Collaborative G. (2018) Association of Body Mass Index and Age With Subsequent Breast Cancer Risk in Premenopausal Women. *JAMA Oncology*. 4: e181771-e.
12. Kabat GC, Kim MY, Lee JS, et al. (2017) Metabolic Obesity Phenotypes and Risk of Breast Cancer in Postmenopausal Women. *Cancer Epidemiol Biomarkers Prev*. 26: 1730-5.
13. Picon-Ruiz M, Morata-Tarifa C, Valle-Goffin JJ, Friedman ER, Slingerland JM. (2017) Obesity and adverse breast cancer risk and outcome: Mechanistic insights and strategies for intervention. *CA: a cancer journal for clinicians*. 67: 378-97.
14. Rose DP, Vona-Davis L. (2010) Interaction between menopausal status and obesity in affecting breast cancer risk. *Maturitas*. 66: 33-8.

15. Althuis MD, Brogan DD, Coates RJ, et al. (2003) Breast cancers among very young premenopausal women (United States). *Cancer Causes & Control*. 14: 151-60.
16. Tavani A, Gallus S, La Vecchia C, et al. (1999) Risk factors for breast cancer in women under 40 years. *European Journal of Cancer*. 35: 1361-7.
17. Dummer TJB, Awadalla P, Boileau C, et al. (2018) The Canadian Partnership for Tomorrow Project: a pan-Canadian platform for research on chronic disease prevention. *Canadian Medical Association Journal*. 190: E710.
18. Amadou A, Hainaut P, Romieu I. (2013) Role of obesity in the risk of breast cancer: lessons from anthropometry. *J Oncol*. 2013: 906495-.
19. White AJ, Nichols HB, Bradshaw PT, Sandler DP. (2015) Overall and central adiposity and breast cancer risk in the sister study. *Cancer*. 121: 3700-8.
20. Huang Z, Willett WC, Colditz GA, et al. (1999) Waist Circumference, Waist: Hip Ratio, and Risk of Breast Cancer in the Nurses' Health Study. *American Journal of Epidemiology*. 150: 1316-24.
21. Lahmann PH, Hoffmann K, Allen N, et al. (2004) Body size and breast cancer risk: Findings from the European prospective investigation into cancer and nutrition (EPIC). *International Journal of Cancer*. 111: 762-71.
22. Harris HR, Willett WC, Terry KL, Michels KB. (2010) Body Fat Distribution and Risk of Premenopausal Breast Cancer in the Nurses' Health Study II. *JNCI: Journal of the National Cancer Institute*. 103: 273-8.
23. Fagherazzi G, Chabbert-Buffet N, Fabre A, et al. (2012) Hip circumference is associated with the risk of premenopausal ER-/PR- breast cancer. *International Journal of Obesity*. 36: 431-9.
24. Carney PL. (2010) Obesity and reproductive hormone levels in the transition to menopause. *Menopause*. 17.
25. Andrew GR, Michelle H, Anthony H. (2006) Insulin-like growth factor (IGF)-I, IGF binding protein-3, and breast cancer risk: eight years on. *Endocrine-Related Cancer*. 13: 273-8.
26. Collaborative Group on Hormonal Factors in Breast Cancer. (2001) Familial breast cancer: collaborative reanalysis of individual data from 52 epidemiological studies including 58 209 women with breast cancer and 101 986 women without the disease. *The Lancet*. 358: 1389-99.
27. Anders CK, Johnson R, Litton J, Phillips M, Bleyer A. (2009) Breast cancer before age 40 years. *Semin Oncol*. 36: 237-49.
28. Peto J, Mack TM. (2000) High constant incidence in twins and other relatives of women with breast cancer. *Nature Genetics*. 26: 411-4.
29. Ginsburg OM, Akbari MR, Aziz Z, et al. (2009) The prevalence of germ-line TP53 mutations in women diagnosed with breast cancer before age 30. *Familial Cancer*. 8: 563.
30. Canadian Cancer Society. (2020) Getting screened for breast cancer.
31. Canadian Task Force on Preventive Health Care. (2020) Breast cancer tools and resources.

32. NCCN Clinical Practice Guidelines In Oncology. (2019) Breast Cancer Screening and Diagnosis. Nation Comprehensive Cancer Network.
33. Klarenbach S, Sims-Jones N, Lewin G, et al. (2018) Recommendations on screening for breast cancer in women aged 40–74 years who are not at increased risk for breast cancer. *Canadian Medical Association Journal*. 190: E1441.
34. Nelson HD, Zakher B, Cantor A, et al. (2012) Risk factors for breast cancer for women aged 40 to 49 years: a systematic review and meta-analysis. *Ann Intern Med*. 156: 635-48.
35. Albrechtsen G, Heuch I, Hansen S, Kvåle G. (2005) Breast cancer risk by age at birth, time since birth and time intervals between births: exploring interaction effects. *British journal of cancer*. 92: 167-75.
36. Liu Q, Wu J, Lambe M, Hsieh S-F, Ekblom A, Hsieh C-C. (2002) Transient increase in breast cancer risk after giving birth: postpartum period with the highest risk (Sweden). *Cancer Causes & Control*. 13: 299-305.
37. Nichols HB, Schoemaker MJ, Cai J, et al. (2018) Breast Cancer Risk After Recent Childbirth. *Ann Intern Med*. 170: 22-30.
38. Russo J, Balogh GA, Russo IH. (2008) Full-term pregnancy induces a specific genomic signature in the human breast. *Cancer Epidemiol Biomarkers Prev*. 17: 51-66.
39. Meier-Abt F, Milani E, Roloff T, et al. (2013) Parity induces differentiation and reduces Wnt/Notch signaling ratio and proliferation potential of basal stem/progenitor cells isolated from mouse mammary epithelium. *Breast Cancer Res*. 15: R36-R.
40. Lyons TR, Schedin PJ, Borges VF. (2009) Pregnancy and breast cancer: when they collide. *J Mammary Gland Biol Neoplasia*. 14: 87-98.
41. Akinyemiju TF, Pisu M, Waterbor JW, Altekruse SF. (2015) Socioeconomic status and incidence of breast cancer by hormone receptor subtype. *Springerplus*. 4: 508.
42. Heck KE, Pamuk ER. (1997) Explaining the Relation Between Education and Postmenopausal Breast Cancer. *American Journal of Epidemiology*. 145: 366-72.
43. Robert SA, Strombom I, Trentham-Dietz A, et al. (2004) Socioeconomic Risk Factors for Breast Cancer: Distinguishing Individual- and Community-Level Effects. *Epidemiology*. 15.
44. Tweed EJ, Allardice GM, McLoone P, Morrison DS. (2018) Socio-economic inequalities in the incidence of four common cancers: a population-based registry study. *Public health*. 154: 1-10.
45. Akinyemiju TF, Demb J, Izano MA, et al. (2018) The association of early life socioeconomic position on breast cancer incidence and mortality: a systematic review. *Int J Public Health*. 63: 787-97.
46. Lundqvist A, Andersson E, Ahlberg I, Nilbert M, Gerdtham U. (2016) Socioeconomic inequalities in breast cancer incidence and mortality in Europe—a systematic review and meta-analysis. *European Journal of Public Health*. 26: 804-13.
47. Gadeyne S, Deboosere P, Vandenheede H, Neels K. (2012) Does birth history account for educational differences in breast cancer mortality? A comparison of premenopausal and postmenopausal women in Belgium. *International journal of cancer*. 131: 2878-85.

48. Lehrer S, Green S, Rosenzweig KE. (2016) Affluence and Breast Cancer. *Breast J.* 22: 564-7.
49. Borugian MJ, Spinelli JJ, Abanto Z, Xu CL, Wilkins R. (2011) Breast cancer incidence and neighbourhood income. *Component of Statistics Canada Catalogue no. 82-003-X.* 22.
50. Goldberg M, Calderon-Margalit R, Paltiel O, et al. (2015) Socioeconomic disparities in breast cancer incidence and survival among parous women: findings from a population-based cohort, 1964-2008. *BMC Cancer.* 15: 921.
51. Menvielle G, Kunst AE, van Gils CH, et al. (2010) The Contribution of Risk Factors to the Higher Incidence of Invasive and In Situ Breast Cancers in Women With Higher Levels of Education in the European Prospective Investigation Into Cancer and Nutrition. *American Journal of Epidemiology.* 173: 26-37.
52. de Kok IMCM, van Lenthe FJ, Avendano M, Louwman M, Coebergh J-WW, Mackenbach JP. (2008) Childhood social class and cancer incidence: Results of the globe study. *Social Science & Medicine.* 66: 1131-9.
53. Gerber B, Müller H, Reimer T, Krause A, Friese K. (2003) Nutrition and lifestyle factors on the risk of developing breast cancer. *Breast Cancer Res Treat.* 79: 265-76.
54. Møller H, Tønnesen H. (1997) Alcohol drinking, social class and cancer. *IARC Sci Publ.* 251-63.
55. Potter JD. (1997) Diet and cancer: possible explanations for the higher risk of cancer in the poor. *IARC scientific publications.* 265-83.
56. Berger E, Castagné R, Chadeau-Hyam M, et al. (2019) Multi-cohort study identifies social determinants of systemic inflammation over the life course. *Nat Commun.* 10: 773.
57. Pepys MB, Baltz ML. (1983) Acute phase proteins with special reference to C-reactive protein and related proteins (pentaxins) and serum amyloid A protein. *Adv Immunol.* 34: 141-212.
58. McDade TW, Lindau ST, Wroblewski K. (2011) Predictors of C-reactive protein in the national social life, health, and aging project. *J Gerontol B Psychol Sci Soc Sci.* 66: 129-36.
59. Kershaw KN, Mezuk B, Abdou CM, Rafferty JA, Jackson JS. (2010) Socioeconomic position, health behaviors, and C-reactive protein: a moderated-mediation analysis. *Health Psychol.* 29: 307-16.
60. Guo L, Liu S, Zhang S, et al. (2015) C-reactive protein and risk of breast cancer: A systematic review and meta-analysis. *Sci Rep.* 5: 10508-.
61. Asegaonkar SB, Asegaonkar BN, Takalkar UV, Advani S, Thorat AP. (2015) C-Reactive Protein and Breast Cancer: New Insights from Old Molecule. *Int J Breast Cancer.* 2015: 145647.
62. Robson PJ, Solbak NM, Haig TR, et al. (2016) Design, methods and demographics from phase I of Alberta's Tomorrow Project cohort: a prospective cohort profile. *CMAJ Open.* 4: E515-E27.