BJR

https://doi.org/10.1259/bjr.20211388

Received: 17 December 2021 | Revised: 12 May 2022 | Accepted: 14 June 2022 | Published online: 07 July 2022

Cite this article as:
Epp J, Rajapakshe R. Breast cancer risk predictions by birth cohort and ethnicity in a population-based screening mammography program. *Br J Radiol* (2022) 10.1259/bjr.20211388.

FULL PAPER

Breast cancer risk predictions by birth cohort and ethnicity in a population-based screening mammography program

[1]JOYCE EPP and [1,2]RASIKA RAJAPAKSHE

[1]BC Cancer – Kelowna, Kelowna, Canada
[2]University of British Columbia Okanagan, Kelowna BC, Canada

Address correspondence to: Dr Rasika Rajapakshe
E-mail: *rrajapak@bccancer.bc.ca*

Objectives: To examine whether birth cohorts affect the risk of breast cancer for East Asian, First Nations, African, South Asian and Caucasian ethnicities in British Columbia (BC).
Methods: We used Cox PH models adjusted for well-known risk factors, such as age, breast density, mammographic features on false positives, and family history, to examine risk of breast cancer among East Asian, First Nations, African and South Asian ethnicities, relative to Caucasian, across three birth cohorts.
Results: There were 813,280 participants and 11,166 *in situ* and invasive breast cancer diagnoses. East Asians screened in BC were found to have a lower risk of breast cancer in the birth cohort born pre-1946 compared to Caucasian, but there was no statistically significant decrease for East Asians born after 1946. First Nations had an increased risk of breast cancer compared with Caucasian for all birth cohorts ranging from 1.1 to 2.0x the risk, which was statistically significant for those born after 1965. South Asians showed a statistically significant decrease in risk ranging from 0.58 to 0.81x lower compared with Caucasians for all birth cohorts.
Conclusion: Risk of breast cancer for South Asians living in BC was found to be lower than Caucasians for each birth cohort examined, while East Asians had a comparable risk of breast cancer, First Nations had a consistently higher risk than Caucasians.
Advances in knowledge: When accounting for birth cohort, compared to Caucasians, South Asians have a decreased risk, First Nations have an increased risk, and East Asians have a similar risk of breast cancer.

INTRODUCTION

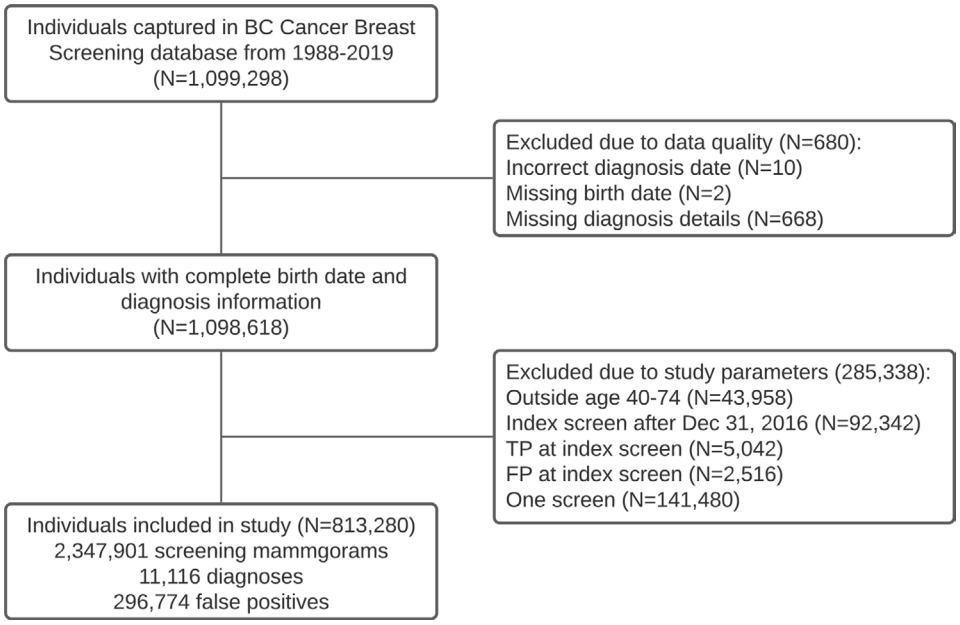
Recent work has shown that predictions of breast cancer risks may be influenced by birth cohort.[1,2] Between birth cohorts, varying lifestyle factors, such as education, diet, and exercise may be present, which may have an effect on breast cancer risk factors.[3,4] Further, the effect of birth cohort may not be consistent across ethnicities.[5] While a comparison of breast cancer risks has been made between Asian countries and the United States while accounting for birth cohort,[1] to our knowledge, the impact of different birth cohorts on breast cancer risk has not been examined with respect to the diverse ethnicities in western countries. Therefore, while adjusting for well-known risk factors, including age, breast density, mammographic features on false positives, and family history, we fit a traditional proportional hazards model, a model stratified by birth cohort, and separate models for each birth cohort, to investigate the changes in risk among East Asian, South Asian, First Nations, African and Caucasian ethnicities in western Canada brought about by birth cohort.

METHODS AND MATERIALS

Participants and study design

This retrospective study uses routinely collected data from the BC Cancer Breast Screening Program, Canada's first population-based screening mammography program. The program provides screening mammograms to females across the province of British Columbia (BC) through 36 fixed and three mobile screening sites.[6,7] The BC Breast Screening program regularly meets national benchmarks for cancer detection and has some of the best survival outcomes in Canada for those who do get breast cancer.[7] Asymptomatic females aged 40–74 with no personal history of breast cancer are eligible to self-refer to screening mammography subject to the guidelines of the program. The main risk factors currently used by BC Cancer Breast Screening to stratify eligible patients are age, family history, and the presence of a high-risk pathologic gene variant, such as BRCA1 and BRCA2.[7] While the information is provided to participants, breast density

Figure 1. Participants excluded from the study. TP = true positive. FP = false positive.



is not used to stratify patients and no risk assessment tool is currently used.

Ethics approval was obtained from the UBC-BC Cancer Research Ethics Board (H18-00172). Informed participant consent was waived by the REB. Data were provided to us de-identified by BC Cancer Breast Screening with a randomly assigned numeric subject identifier. The data contained risk factors, screening, and diagnosis information routinely collected for individuals who attended screening mammography from 1988 through 2016. Participants were followed for 5 years, until invasive or *in situ* breast cancer diagnosis, until their last screen if there was no diagnosis, or until study end, whichever came first. Participants were excluded if they were outside of the age range of 40–74, had a true positive or false negative on their index screen, or had missing birth date or diagnosis information. The analysis was done on the remaining cohort of 813,280 participants, 11,166 invasive and *in situ* breast cancer diagnoses, and 2,347,901 screening mammograms (Figure 1).

The outcome was the occurrence of an individual's first *in situ* or invasive breast cancer diagnosis. Education, ethnicity, age at menarche, age at first child, and number of children were self-reported at the first screening mammogram appointment. Mother and/or sister with a history of breast cancer, estrogen use at the time of the screening appointment, and previous breast biopsy were collected at each screening mammogram appointment with a self-administered questionnaire. Mammographic features were included as a categorical variable indicating the presence of one feature or combinations of features. Breast density, where available, was coded as a binary variable, representing $< 50\%$ and $\geq 50\%$ breast density. The breast density and mammographic features were ascertained and recorded by trained and credentialed radiologists and rely completely on the radiologist's interpretation of the mammogram. BC Cancer radiologists must have a minimum of 2 years' experience with a minimum of 2,500 mammogram interpretations and must perform appropriately on a standardized test before being permitted to interpret screens. They are required to maintain an annual volume of 1,500 screen interpretations and are credentialed with the Screening program.[8] Radiologists use the Breast Imaging Reporting and Data System (BIRADS) lexicon to interpret screening mammograms.[9,10]

Statistical methods

All analyses were done using SAS and R.[11,12] Data were checked for missing at random and missing data for family history were given the baseline value of "No". Participants without an ethnicity recorded were given the ethnicity "Other". Missing mammographic features for participants with an abnormal screen were coded as "Unknown". The remaining missing data were imputed using SAS software.[13] Variable selection was performed using backwards selection with a p-value < 0.05.[14,15] The data were split into a randomly selected training and testing set, where 70% was used for training and 30% for testing. To create the training and testing sets, we used a built-in function in R, which takes a simple random sample without replacement of a specified size from the population and assigns those individuals to the training group, with the remainder placed into the testing group. Loss to follow-up was addressed by following non-diagnosed individuals until the last known screen and potential bias was addressed by removing individuals with a true positive or false negative on the first screen.

A Cox proportional hazards model and time-varying Cox model were developed on the training set and validated on the test set.[16,17] The predictors in the Cox proportional hazards model were measured at the index screening mammogram. In the time-varying model, all the predictors were coded in the same way except for breast density and mammographic features,

Table 1. Description of the data

Risk factor	Missing n, (%)	Average (SD) or count (%)
Total *n* = 813,280		
Sex (Female)	0	813,280 (100)
Age (years)	0	51.8 (9.3)
Ethnicity	305,596 (37.5)	
First Nations		9,911 (1.2)
African		1,794 (0.2)
East Asian		85,512 (10.5)
Caucasian		382,942 (47.0)
South Asian		27,525 (3.4)
Education	60,897 (7.5)	
No high school		49,542 (6.1)
High school or some high school		295,497 (36.3)
College grad or some college		407,344 (50.1)
Number of children	2,617 (0.3)	1.9 (1.5)
Mother	0	
No breast cancer		753,566 (92.7)
Previous breast cancer		59,714 (7.3)
Sister	0	
No breast cancer		779,616 (95.9)
Previous breast cancer		33,664 (4.1)
Current estrogen use	6,588 (0.8)	
No		702,591 (86.4)
Yes		104,101 (12.8)
Ever had breast biopsy	5,010 (0.6)	
No		730,691 (89.8)
Yes		77,579 (9.5)
Breast density	143,807 (17.7)	
<50%		433,171 (53.3)
>=50%		236,302 (29.1)
Mammographic feature	95,275 (11.7)	
Normal screen		553,486 (68.1)
Asymmetry only		82,510 (10.1)
Calcification only		23,823 (2.9)
Mass only		33,916 (4.2)
Distortion only		9,941 (1.2)
Distortion and asymmetry		5,350 (0.7)
Mass and asymmetry		3,135 (0.4)
Asymmetry and calcification		2,685 (0.3)
Mass and calcification		1,547 (0.2)
Mass and distortion		582 (0.1)
Three or more		578 (0.1)
Distortion and calcification		452 (0.1)

which were treated as time-varying, meaning that a different mammographic feature or breast density value on a subsequent mammogram would be accounted for in the model.

To assess model fit, we checked for proportional hazards, influential observations, and non-linearity. Proportional hazards were checked using a graphical check of the Schoenfeld residuals and no strong violations were found.[18,19] Influential observations were assessed using dfbetas and non-linearity in the only numeric variable, age, was assessed using Martingale residuals.[19–21]

Five-year absolute risk, concordance and calibration were calculated in R on the test set using the models developed on the training set.[12,15,19,21] Five-year absolute risk is the probability of getting a breast cancer diagnosis within 5 years. Using the 5-year risk thresholds identified in the IBIS model,[22] we identified the proportions at general and high risk for ages 40–49 and 50–70. To compare the increase in risk in the high risk group to the general risk group, stratified risk was calculated by dividing the mean risk in the high risk group by the mean risk in the general risk group. Concordance is a measure of discrimination where a concordance of 1 indicates perfect discrimination, or individuals without a diagnosis had consistently lower risk scores than individuals who are diagnosed. A value of 0.5 indicates that the model is no better than randomly guessing at distinguishing between who will get a diagnosis and who will not.[23,24] Calibration is a 'goodness-of-fit' measure which gives the ratio of the observed number of cases to predicted number of cases.[25,26] A well-calibrated model will have a ratio close to 1.

We then adjusted our Cox PH model by birth cohort and fit separate models for each birth cohort. Birth cohorts used were: participants born prior to 1946, born between 1946 and 1965, and born after 1965. These birth cohorts were chosen to capture the three large generations in Canada marked by the baby boom within 1946–1965. We wanted to capture the changes in lifestyle that occurred during these timeframes, as indicated by changing fertility rates.[27]

RESULTS

We used 813,280 participants in the analysis, with a median age of 50 (IQR: 44–58) within the age range of 40–75. (Figure 1). During the median follow-up time of 4.97 and 5 years for participants with and without a false positive, respectively, we had a mean of three screening mammograms per participant and 11,166 diagnoses overall. Overall cancer detection rate was 4.7 per 1000, which both meets and exceeds the Canadian threshold of 3 per 1000 and 1 per 1000 for invasive and *in situ*, respectively.[6,7,9,10] Of the 2,347,901 screening mammograms, 296,774 (12.6%) were false positives with features as described in Table 1. The mammographic feature recorded the most often on false positives was asymmetry and the combination of architectural distortion with calcifications occurred least often. With respect to ethnicity, 47% reported one of the ethnicities that were used to create the Caucasian group, 10% reported East Asian, 3% South Asian and only 1.2 and 0.2% First Nations and African, respectively. The remaining 38% of the participants did not have an ethnicity recorded, and are in the group 'Other'. Table 1 gives a description of the baseline data values. We confirmed that breast density decreases with age and is stratified by ethnicity (Figure 2). Average breast density was calculated by assigning < 50% breast density a numeric value of 1 and ≥ 50% breast density a value of 2 and averaging these values for the 5 year age groups. We also confirmed that number of children is affected by birth cohort (Figure 3). Figure 3 shows the changing trends in number of children based on year of birth and is an indication of lifestyle differences between the birth cohorts.

The concordance for training and test set were 0.638 (95% CI 0.635, 0.641) and 0.627 (95% CI 0.621, 0.633), respectively, in the Cox PH model (Figure 4).[12,24] The calibration was assessed at 0.96 for 5 year risk.[12,15,26] The overall 5 year mean absolute risk

Figure 2. Average breast density for 5-year age groups at index screen stratified by ethnicity.

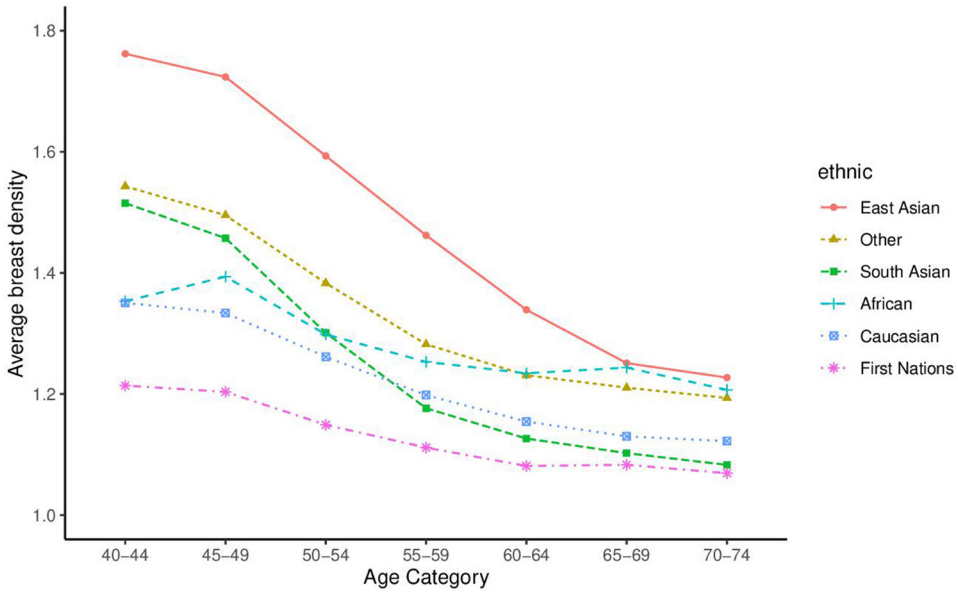
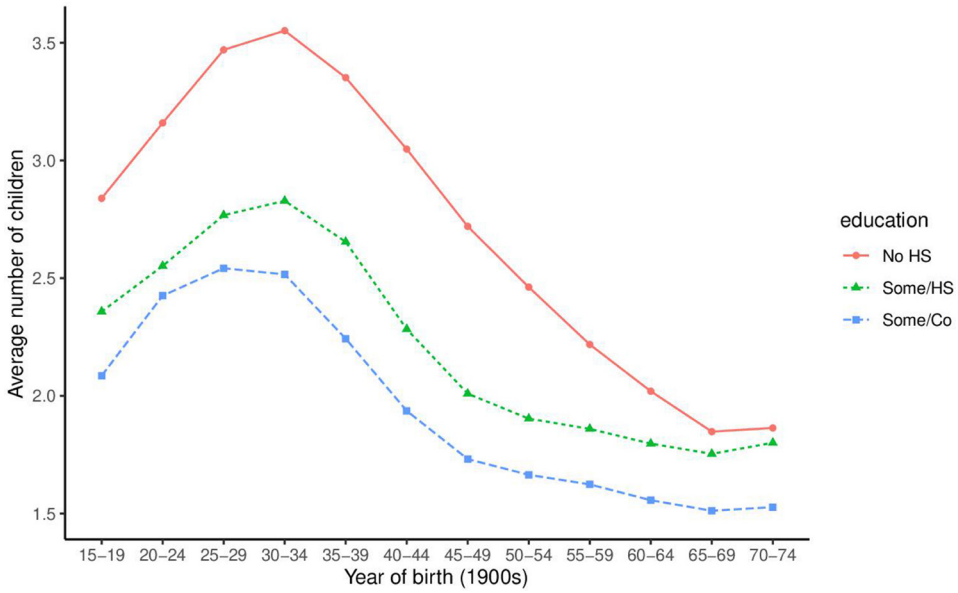


Figure 3. Number of children stratified by self-reported education level at index screen. HS = high school. CO = college.



was 1.61%, while the mean absolute risk for those with no diagnosis and those with a diagnosis was 1.6 and 2.0%, respectively ($p < 0.001$).[12,28–30]

Table 2 shows that most of the females were predicted by the model to fall within the general risk group, however over 15% of females between 50 and 70 were grouped into a high risk group. For females between 50 and 70, the 5 year risk of breast cancer for the general risk group was 1.6%, while the high risk group was 1.9x that amount at 3%.

Table 3 shows the Cox model built on the full data and adjusted for birth cohort. The basic Cox PH model and model with time-varying mammographic features are also provided for comparison. First Nations participants when compared to Caucasian participants have an increase in risk in these models (Table 3). However, in Figure 1 it was shown that First Nations had the lowest breast density. South Asian participants have a decrease in risk compared with Caucasians. Both African and East Asian participants do not show a strong statistically different risk relative to Caucasians in the models.

We grouped individuals by three birth cohorts which characterize the three largest generations in Canada (born prior to 1946, between 1946 and 1965, and after 1965) to investigate whether they had an effect on the risks seen by different ethnic groups, and fit a separate model for each birth cohort (Table 4). African participants were found to be at a statistically significant higher risk of breast cancer than Caucasians for those born after 1965. Females of Caribbean descent were included in the African

Figure 4. Discrimination for the basic Cox PH model separated by diagnosis status.

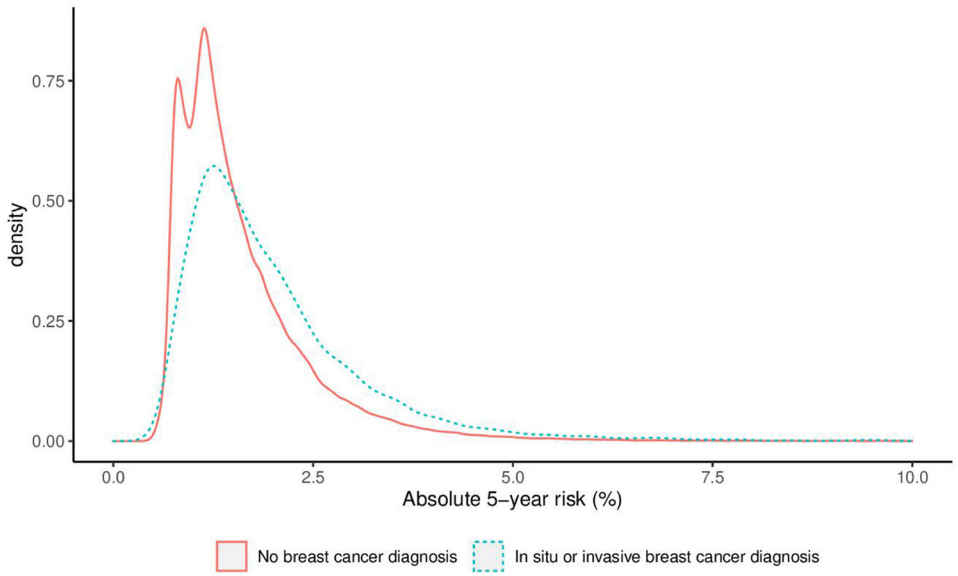


Table 2. Predictive performance of the proportional hazards model

	Cox proportional hazards model absolute risk predictions				
	Risk group	**Risk range**	**% at risk**	**Mean risk**	**Stratified risk**
40–49 y	General	0–1.5	85.6	1.0	1.0 (reference)
	High	1.5–2.2	11.0	1.8	1.7
	Above high	>2.2	3.3	2.7	2.7
50–70 y	General	0–2.5	81.5	1.6	1.0 (reference)
	High	2.5–4	15.0	3.0	1.9
	Above high	>4	3.5	5.2	3.3

ethnicity as these data were not available from the Screening program.[31] South Asian participants were at a consistently lower risk than Caucasian, regardless of the birth cohort. We examined whether risk factors affected by lifestyle would explain these results, but our analysis did not reveal a difference in trends for South Asian birth cohorts compared to other ethnicities for: age at first childbirth, number of children, level of education, age at menarche, or age at menopause. Among all ethnicities we examined, we found that the age at first delivery is increasing, number of children is decreasing, education level is increasing, and age at menarche is slightly decreasing (Supplementary Appendix A). As these common risk factors did not provide an indication as to being the cause of our results, since each ethnicity had a similar trend, further investigation into why variation of risk exists between ethnicities for the birth cohorts examined is suggested.

DISCUSSION

To investigate risk of breast cancer for different birth cohorts among East Asian, First Nations, African and South Asian ethnicities in British Columbia, we examined several models to predict risk of breast cancer for individuals attending routine breast screening in British Columbia. We included mammographic features on false positives as they have been shown to increase the risk of breast cancer, but have not yet been examined in conjunction with estimating risk of breast cancer among ethnic groups.[32] Treating the mammographic features as time-varying did not affect the risk for different ethnicities. Adjusting the model for birth cohort also did not affect the risk for different ethnicities. Building separate models for each birth cohort found that, for participants born prior to 1946, First Nations had a 1.1x higher risk than Caucasians, East Asians had 0.73x the risk, and South Asians had 0.66x the risk of Caucasians. For participants born within the baby boom from 1946 to 1965, First Nations had 1.2x the risk of Caucasians, East Asians had a comparable risk to Caucasians, and South Asians had 0.91x the risk of Caucasians. For participants born after 1965, First Nations had 2x the risk of Caucasians, East Asians had a comparable risk to Caucasians, and South Asians had 0.58x the risk of Caucasians.

We found that the relative risk of 0.74 for East Asians compared to Caucasians found by Hoegg[33] was corroborated with a relative risk of 0.73 for East Asians in only the cohort born prior to 1946.[12,19] A previous study from Sung et al has similarly found that the incidence rates of East Asian and Western individuals are converging.[1] They suggest that the previously seen protective effect of Asian ethnicity was likely due to lifestyle. Asians living in BC may be subject to the same change as they adapt a more western lifestyle. First Nations also have a consistently higher risk, regardless of birth cohort, and this risk may be increasing, as seen by the slightly higher risk in the birth cohort post-1965 (Table 4).

The African ethnic subgroup showed a relative risk 1.2x that of Caucasians for females born prior to 1946 and a risk 3.4x times that of Caucasians for females born after 1965, although only the latter was statistically significant (Table 4). Given the known aggressiveness of breast cancer in the African population, an increased risk of breast cancer compared to Caucasians is concerning and may warrant additional supplemental screening procedures for this subgroup.[34,35] For the group born during the baby boom between 1946 and 1965, the risk of breast cancer was 0.53x that of Caucasians, which was not statistically significant. This apparent protective effect may have a number of causes, such as, a result of fewer early-stage cancers being detected by screening in African females, skew by our small sample size, or unknown lifestyle factors present for females in this birth cohort.[35]

It is well known that a high breast density is a strong risk factor for breast cancer. We found First Nations to have the lowest mammographic density, however, they were found to have a risk 2x that of Caucasians for participants born post-1965. As such, this may suggest that a high breast density should be taken into account with ethnicity and other factors.

This research is strengthened by the large dataset with many routinely collected demographic and lifestyle risk factors included. However, our data are limited by participant recall of their own personal risk factors, as participant recall may be biased. As well, mammographic information is dependent on the radiologists' interpretation. Due to inconsistent recording of mammographic breast density in the Screening program, mammographic density was missing in 17% of cases. Thus, a limitation to this study is that we relied on imputation for these missing mammographic density cases. 38% of the participants did not self-report their ethnicity to the Screening program. It is expected that the missing ethnicities would follow a similar distribution as the observed ethnicities, since there is less than a 3% difference between the Screening Program attendance ethnic proportions and the ethnic proportions in the province of British

Table 3. Models

	HR (95% CI)		
Risk Factor	**Cox PH model**	**Time-varying Cox PH model**	**Cox PH model adjusted for birth cohort**
	***N* = 569,296**	***N* = 569,296**	***N* = 813,280**
Mammographic feature			
Normal screen (ref)	1.00	1.00	1.00
Asymmetry only	0.91 (0.83–0.99)	1.38 (1.20–1.58)	0.91 (0.85–0.98)
Calcification only	1.97 (1.78–2.17)	2.74 (2.37–3.17)	1.97 (1.81–2.15)
Mass only	1.11 (0.99–1.24)	1.83 (1.56–2.15)	1.11 (1.00–1.22)
Distortion only	1.06 (0.87–1.29)	1.68 (1.20–2.34)	1.08 (0.91–1.27)
Distortion and asymmetry	1.06 (0.79–1.43)	1.48 (0.88–2.51)	1.07 (0.83–1.37)
Mass and asymmetry	1.30 (0.92–1.85)	1.60 (0.93–2.76)	1.51 (1.15–1.99)
Asymmetry and calcification	1.70 (1.23–2.36)	2.62 (1.67–4.07)	1.82 (1.39–2.38)
Mass and calcification	1.58 (1.04–2.40)	1.93 (1.04–3.61)	1.48 (1.02–2.14)
Mass and distortion	2.12 (1.14–3.94)	3.78 (1.70–8.42)	2.53 (1.57–4.07)
Three or more	1.90 (0.99–3.67)	4.26 (1.91–9.48)	1.97 (1.14–3.40)
Distortion and calcification	2.64 (1.46–4.77)	3.52 (1.47–8.47)	2.76 (1.69–4.52)
Unknown	1.21 (1.13–1.28)	1.78 (1.64–1.94)	1.18 (1.12–1.24)
Breast density			
<50% (ref)	1.00	1.00	1.00
≥50%	1.49 (1.42–1.57)	1.71 (1.62–1.80)	1.48 (1.42–1.54)
Age	1.03 (1.03–1.04)	1.04 (1.03–1.04)	1.03 (1.03–1.03)
Ethnicity			
Caucasian (ref)	1.00	1.00	1.00
First Nations	1.24 (1.02–1.50)	1.21 (1.01–1.50)	1.24 (1.05–1.46)
African	1.05 (0.66–1.67)	0.96 (0.57–1.62)	1.00 (0.67–1.50)
East Asian	0.92 (0.85–1.00)	0.86 (0.79–0.94)	0.95 (0.89–1.01)
Other	0.98 (0.93–1.03)	0.96 (0.91–1.02)	1.00 (0.96–1.05)
South Asian	0.76 (0.66–0.89)	0.74 (0.63–0.87)	0.74 (0.65–0.84)
Education			
No high school (ref)	1.00	1.00	1.00
High school or some high school	1.21 (1.10–1.32)	1.20 (1.08–1.32)	1.25 (1.15–1.35)
College grad or some college	1.29 (1.17–1.42)	1.26 (1.14–1.39)	1.31 (1.21–1.42)
Mother			
No breast cancer (ref)	1.00	1.00	1.00
Previous breast cancer	1.64 (1.53–1.76)	1.68 (1.55–1.80)	1.61 (1.52–1.71)
Sister			
No breast cancer (ref)	1.00	1.00	1.00
Previous breast cancer	1.41 (1.29–1.54)	1.41 (1.28–1.55)	1.48 (1.37–1.59)
Current estrogen use			
No (ref)	1.00	1.00	1.00
Yes	1.17 (1.10–1.24)	1.16 (1.09–1.24)	1.18 (1.12–1.24)

(Continued)

Table 3. (Continued)

	HR (95% CI)		
Ever had breast biopsy			
No (ref)	1.00	1.00	1.00
Yes	1.41 (1.33–1.50)	1.40 (1.31–1.49)	1.40 (1.33–1.47)

Columbia (Supplementary Appendix A Table 1Supplementary Appendix A Table 1).[36] Having this information would decrease the confidence intervals around the risk parameter estimates.

The retrospective design is a weakness as several important risk factors were not available for consideration during the period under study, such as body mass index, genetic mutations, prior mantle radiation, and a previous diagnosis of a high-risk lesion. However, we have included several important risk factors that were available from the Screening Program, such as age, breast density, and family history.

Due to inclusion of a high number of baseline screens in this study, we observed a higher than average number of false positives compared to the Canadian standards (12% *vs* 7–8%). We also observed an increase in risk of 1.7x for females with ≥ 50% breast density compared to < 50% breast density, while the literature reports an increased risk of 2x higher. An explanation for this could be due to the imputation of the missing breast density information, which is a limitation to this study.

Over 80% of participants attending screening in British Columbia born after 1975 have at least some post-secondary education. Thus, including education with 'no high school' as the reference may not be relevant in a few years. Future work might focus on more effective predictors using a more comprehensive measure of lifestyle. We did not observe a protective effect for East Asian ethnicity as has been previously found,[33] except for in the birth cohort born prior to 1946. Previous research has suggested that that the protective effect of East Asian ethnicity is likely due to lifestyle.[1] African and Caribbean participants were found to have an increased risk in the birth cohorts born prior to 1946 and after 1965, but a protective effect was found

Table 4. Risk of breast cancer for First Nations, East Asian, and South Asian ethnicities relative to Caucasian for each birth cohort

Ethnicity	n	HR (95% CI)	*p*-value
	Born prior to 1946 (n = 264,014)		
Caucasian	119,442	1.00 (ref)	
First Nations	2,659	1.14 (0.86–1.51)	0.35
African	456	1.22 (0.67–2.21)	0.51
East Asian	20,263	0.73 (0.64–0.83)	<0.001
Other	114,482	0.98 (0.93–1.04)	0.55
South Asian	6,712	0.66 (0.53–0.83)	<0.001
	Born between 1946 and 1965 (n = 438,128)		
Caucasian	236,384	1.00 (ref)	
First Nations	6,420	1.20 (0.96–1.49)	0.10
African	1,119	0.53 (0.25–1.11)	0.09
East Asian	56,464	1.05 (0.97–1.14)	0.25
Other	120,486	1.00 (0.94–1.07)	0.95
South Asian	17,255	0.81 (0.69–0.95)	0.01
	Born after 1965 (n = 111,138)		
Caucasian	27,116	1.00 (ref)	
First Nations	832	2.01 (1.17–3.45)	0.01
African	219	3.39 (1.51–7.62)	0.003
East Asian	8,785	1.17 (0.91–1.49)	0.22
Other	70,628	0.98 (0.83–1.16)	0.80
South Asian	3,558	0.58 (0.36–0.94)	0.03

Note: Participants of Caribbean descent are included in the African ethnicity

for females born in 1946–1965. Further work should be done to investigate and corroborate these findings with a larger African subgroup, especially given the aggressiveness of breast cancer for African females. Unfortunately, our study cohort did not contain large enough females with African or Caribbean descent to make any conclusions. We found that First Nations participants have a consistently higher risk compared to Caucasian participants and South Asian participants have a consistently lower risk for all birth cohorts examined. One potential hypothesis could be that South Asian participants and their descendants are maintaining their diet and lifestyle compared to East Asian participants. Validation of these results is required and further investigation should be made into why South Asian participants have a consistently lower risk than Caucasian, despite the similar changes in risk factors. If the difference in risk is due to lifestyle or diet, then more useful predictors that are a more comprehensive measure of lifestyle should be sought after and which aspects of lifestyle should be investigated further. In conclusion, we find that further studies into the differences in risk by ethnicity and birth cohort may provide insight into breast cancer risk.

ACKNOWLEDGMENT

The authors wish to acknowledge Amanda Wanner for her help with the manuscript preparation and submission; Colin Mar and Janette Sam for their comments on the manuscript; Janette Sam for her assistance on data access; Mikael Hartman for his advice on cohort analysis, and; Jeff Andrews for his advice on statistical analysis.

FUNDING

The authors have no funding or conflicts of interest to report.

REFERENCES

1. Sung H, Rosenberg PS, Chen W-Q, Hartman M, Lim W-Y, Chia KS, et al. Female breast cancer incidence among asian and western populations: more similar than expected. *J Natl Cancer Inst* 2015; **107**(7): djv107. https://doi.org/10.1093/jnci/djv107
2. Epp J. Development of a prognostication model for breast cancer in British Columbia. Master's thesis. University of British Columbia. 2020. Available from: http://hdl.handle.net/2429/75136
3. Keyes KM, Utz RL, Robinson W, Li G. What is a cohort effect? comparison of three statistical methods for modeling cohort effects in obesity prevalence in the united states, 1971-2006. *Soc Sci Med* 2010; **70**: 1100–1108. https://doi.org/10.1016/j.socscimed.2009.12.018
4. Dong JY, Qin LQ. Education level and breast cancer incidence: a meta-analysis of cohort studies. *Menopause* 2020; **27**: 113–18. https://doi.org/10.1097/GME.0000000000001425
5. Liu L, Zhang J, Wu AH, Pike MC, Deapen D. Invasive breast cancer incidence trends by detailed race/ethnicity and age. *Int J Cancer* 2012; **130**: 395–404. https://doi.org/10.1002/ijc.26004
6. Cancer BC. BC Cancer Breast Screening 2019 Program Results. Vancouver BC. 2020. Available from: http://www.bccancer.bc.ca/screening/Documents/Breast-Screening-Program-Report-2019.pdf
7. Agency BC. Celebrating 25 Years of Breast Screening in BC. Vancouver, BC. 2015. Available from: http://www.bccancer.bc.ca/screening/Documents/SMP_Report-25thAnniversaryReport.pdf
8. Burhenne LJ, Burhenne HJ, Kan L. Quality-oriented mass mammography screening. *Radiology* 1995; **194**: 185–88. https://doi.org/10.1148/radiology.194.1.7997549
9. BC Cancer. BC Cancer Breast Screening 2017 Program Results. Vancouver BC. 2018. Available from: http://www.bccancer.bc.ca/screening/Documents/Breast_AnnualReport2018.pdf
10. BC Cancer. BC Cancer Breast Screening 2019 Program Results. Vancouver BC. 2020. Available from: http://www.bccancer.bc.ca/screening/Documents/Breast-Screening-Program-Report-2019.pdf
11. SAS Institute Inc. *SAS University Edition. SAS Institute Inc*. Cary, NC, USA; 2019.
12. R Core Team. *R: A Language and Environment for Statistical Computing*. Vienna, Austria: R Foundation for Statistical Computing; 2019.
13. SAS Institute Inc. *SAS Institute Inc. SAS/STAT 14.1 User's Guide*. Gary, NC, USA; 2015.
14. Sakamoto Y, Ishiguro M, Kitagawa G. *Akaike information criterion statistics*. Dordrecht, The Netherlands: D. Reidel; 1986.
15. Mogensen UB, Ishwaran H, Gerds TA. Evaluating random forests for survival analysis using prediction error curves. *J Stat Softw* 2012; **50**: 1–23. https://doi.org/10.18637/jss.v050.i11
16. Cox DR. Regression models and life-tables. *Journal of the Royal Statistical Society: Series B (Methodological)* 1972; **34**: 187–202. https://doi.org/10.1111/j.2517-6161.1972.tb00899.x
17. Fisher LD, Lin DY. Time-dependent covariates in the cox proportional-hazards regression model. *Annu Rev Public Health* 1999; **20**: 145–57. https://doi.org/10.1146/annurev.publhealth.20.1.145
18. Schoenfeld D. Partial residuals for the proportional hazards regression model. *Biometrika* 1982; **69**: 239–41. https://doi.org/10.1093/biomet/69.1.239
19. Therneau TM, Lumley T. Package 'survival. *R Top Doc* 2015; **128**: 28–33.
20. Wu CFJ. Jackknife, bootstrap and other resampling methods in regression analysis. *Ann Statist* 1986; **14**: 1261–95. https://doi.org/10.1214/aos/1176350142
21. Therneau TM, Grambsch PM, Fleming TR. Martingale-based residuals for survival models. *Biometrika* 1990; **77**: 147–60. https://doi.org/10.1093/biomet/77.1.147
22. MacInnis RJ, Knight JA, Chung WK, Milne RL, Whittemore AS, Buchsbaum R, et al. Comparing 5-year and lifetime risks of breast cancer using the prospective family study cohort. *J Natl Cancer Inst* 2021; **113**: 785–91. https://doi.org/10.1093/jnci/djaa178
23. Harrell FE. Rms: regression modeling strategies. *R Package Version 51-31* 2019.
24. Harrell FE, Lee KL, Mark DB. Multivariable prognostic models: issues in developing models, evaluating assumptions and adequacy, and measuring and reducing errors. *Stat Med* 1996; **15**: 361–87. https://doi.org/10.1002/(SICI)1097-0258(19960229)15:4<361::AID-SIM168>3.0.CO;2-4
25. Harrell FE Jr, Lee KL, Califf RM, Pryor DB, Rosati RA. Regression modelling strategies for improved prognostic prediction. *Stat Med* 1984; **3**: 143–52. https://doi.org/10.1002/sim.4780030207
26. van Houwelingen HC. Validation, calibration, revision and combination of prognostic survival models. *Stat Med* 2000; **19**: 3401–15. https://doi.org/10.1002/1097-

0258(20001230)19:24<3401::aid-sim554>3.0.co;2-2
27. Statistics Canada. Generations in Canada. 2018. Available from: https://www12.statcan.gc.ca/census-recensement/2011/as-sa/98-311-x/98-311-x2011003_2-eng.cfm
28. Schoenfeld D. Chi-squared goodness-of-fit tests for the proportional hazards regression model. *Biometrika* 1980; **67**: 145–53. https://doi.org/10.1093/biomet/67.1.145
29. Kruskal WH, Wallis WA. Use of ranks in one-criterion variance analysis. *Journal of the American Statistical Association* 1952; **47**: 583–621. https://doi.org/10.1080/01621459.1952.10483441
30. Steyerberg EW, Vickers AJ, Cook NR, Gerds T, Gonen M, Obuchowski N, et al. Assessing the performance of prediction models: a framework for traditional and novel measures. *Epidemiology* 2010; **21**: 128–38. https://doi.org/10.1097/EDE.0b013e3181c30fb2
31. Hwee J, Bougie E. Do cancer incidence and mortality rates differ among ethnicities in canada? *Health Rep* 2021; **32**: 3–17. https://doi.org/10.25318/82-003-x202100800001-eng
32. Castells X, Torá-Rocamora I, Posso M, Román M, Vernet-Tomas M, Rodríguez-Arana A, et al. Risk of breast cancer in women with false-positive results according to mammographic features. *Radiology* 2016; **280**: 379–86. https://doi.org/10.1148/radiol.2016151174
33. Hoegg T. Statistical modelling of breast cancer risk for British Columbian women. Master's thesis. University of British Columbia. 2013. Available from: http://hdl.handle.net/2429/44569
34. Iqbal J, Ginsburg O, Rochon PA, Sun P, Narod SA. Differences in breast cancer stage at diagnosis and cancer-specific survival by race and ethnicity in the united states. *JAMA* 2015; **313**: 165–73. https://doi.org/10.1001/jama.2014.17322
35. Daly B, Olopade OI. A perfect storm: how tumor biology, genomics, and health care delivery patterns collide to create A racial survival disparity in breast cancer and proposed interventions for change. *CA Cancer J Clin* 2015; **65**: 221–38. https://doi.org/10.3322/caac.21271
36. Statistics Canada. Ottawa: Census Profile. Available from: https://www12.statcan.gc.ca