



Published in final edited form as:

Cancer Res. 2013 December 15; 73(24): 7243–7253. doi:10.1158/0008-5472.CAN-13-2014.

Genetic ancestry and risk of mortality among U.S. Latinas with breast cancer

Laura Fejerman^{1,*}, Donglei Hu², Scott Huntsman², Esther M. John^{3,4}, Mariana C. Stern⁵, Christopher A. Haiman⁶, Eliseo J. Pérez-Stable⁷, and Elad Ziv⁸

¹Department of Medicine, Division of General Internal Medicine, Institute for Human Genetics, Helen Diller Family Comprehensive Cancer Center and Center for Aging in Diverse Communities, UCSF, San Francisco, CA 94158

²Department of Medicine, Division of General Internal Medicine, UCSF, San Francisco, CA 94158

³Cancer Prevention Institute of California, Fremont, CA 94538

⁴Division of Epidemiology, Department of Health Research and Policy, and Stanford Cancer Institute, Stanford University School of Medicine, Stanford, CA 94503

⁵Department of Preventive Medicine, Norris Comprehensive Cancer Center, Keck School of Medicine of USC, Los Angeles, CA 90089

⁶Department of Preventive Medicine, Norris Comprehensive Cancer Center, Keck School of Medicine, University of Southern California. Los Angeles, CA 90033, US

⁷Department of Medicine, Division of General Internal Medicine, Medical Effectiveness Research Center for Diverse Populations, and Helen Diller Family Comprehensive Cancer Center UCSF, San Francisco, CA 94143

⁸Department of Medicine, Division of General Internal Medicine, Institute for Human Genetics and Helen Diller Family Comprehensive Cancer Center, UCSF, San Francisco, CA 94158

Abstract

Multiple studies have reported that Latina women in the U.S. are diagnosed with breast cancer at more advanced stages and have poorer survival than non-Latina White women. However, Latinas are a heterogeneous group with individuals having different proportions of European, Indigenous American and African genetic ancestry. In this study we evaluated the association between genetic ancestry and survival after breast cancer diagnosis among 899 Latina women from the San Francisco Bay Area. Genetic ancestry was estimated from single nucleotide polymorphisms from an Affymetrix 6.0 array and we used Cox proportional hazards models to evaluate the association between genetic ancestry and breast cancer-specific mortality (tests were two-sided). Women were followed for an average of 9 years during which 75 died from breast cancer. Our results showed that Individuals with higher Indigenous American ancestry had increased risk of breast cancer-specific mortality [hazard ratio (HR): 1.57 per 25% increase in Indigenous American ancestry; 95% confidence interval (CI): 1.08–2.29]. Adjustment for demographic factors, tumor

*To whom correspondence should be addressed. Tel: 415-514-4934; Fax: 415-514-34982; laura.fejerman@ucsf.edu.

Conflict of Interest: There are no conflict of interests to disclosed

characteristics, and some treatment information did not explain the observed association [HR: 1.75, 95%CI: 1.12–2.74]. In an analysis in which ancestry was dichotomized, the hazard of mortality showed a two-fold increase when comparing women with <50% Indigenous American ancestry to women with ≥50% [HR: 1.89, 95%CI: 1.10–3.24]. This was also reflected by Kaplan-Meier survival estimates (P for Log-Rank test of 0.003). Overall, results suggest that genetic factors and/or unmeasured differences in treatment or access to care should be further explored to understand and reduce ethnic disparities in breast cancer outcomes.

Keywords

Breast cancer mortality; U.S. Latinas; Genetic ancestry

Introduction

The incidence of breast cancer varies by race/ethnicity, with higher rates in non-Latina White women than African American, Asian/Pacific Islander, Hispanic/Latina or American Indian/Alaskan native women (1, 2). Despite the lower risk of breast cancer among U.S. women from minority populations, their risk of mortality after diagnosis is higher compared to that of non-Latina White women (3–7). In particular, compared to non-Latina Whites, Latina women are diagnosed with more advanced cancer stage and have poorer survival after controlling for prognostic factors (i.e., disease stage, age, tumor histology and treatment) (3, 8). Some studies have suggested that differences in survival may be explained by differences in socioeconomic status and access to care through their effects on disease stage at diagnosis and other tumor characteristics, as well as through direct effects on survival (4–7). However, most studies have not considered the well-known genetic heterogeneity among people within each racial/ethnic group.

Latinos have substantial genetic variation in terms of European, Indigenous American and African ancestry (9–12). In California, the majority of Latinos originates from Mexico and from Central America and has mostly European and Indigenous American ancestry. We have previously demonstrated an association between breast cancer risk and genetic ancestry among Latina women (13–15). In particular, among both U.S. Latinas and Mexican women, we found that higher European ancestry was associated with increased risk, whereas higher Indigenous American ancestry was associated with decreased risk (14, 15). This association is consistent with the epidemiological observation that non-Latina White women in the U.S. have higher incidence of breast cancer compared to both Latina and Native American women.

Previous studies suggested that breast cancer prognosis might have an inherited component (16, 17), and therefore exploring the relationship between genetic ancestry and breast cancer-specific mortality among women of mixed European, Indigenous American and African ancestry could provide important insights. Given that breast cancer mortality among Latinas is higher compared to non-Latina Whites, we hypothesized that Indigenous American ancestry may be associated with increased breast cancer-specific mortality. We tested this hypothesis in a sample of 899 U.S. Latina women with breast cancer. We also

evaluated whether the association between genetic ancestry and breast cancer-specific mortality was explained by differences in socioeconomic status, education, body mass index, family history of breast cancer, place of birth, language use, age at cancer diagnosis, treatment received and tumor characteristics.

Materials and Methods

Samples

The San Francisco Bay Area Breast Cancer Study (SFBCS): This population-based case-control study in Latinas, African-American, and non-Latina White women aged 35–79 years was conducted between 1995 and 2004. A detailed description is provided elsewhere (18, 19). The present study includes samples from 333 Latina women genotyped using the Affymetrix 6.0 array.

Northern California site of the Breast Cancer Family Registry (NC-BCFR): This population-based family study enrolled incident breast cancer cases aged 18–64 years and selected family members. Cases include all those with indicators of increased genetic risk (i.e., cases with early onset disease, bilateral disease, personal history of ovarian cancer, family history of breast or ovarian cancer) and a random sample of cases not meeting these criteria (20). The present study includes 594 unrelated female cases diagnosed from 1995–2003 and that were genotyped using the Affymetrix 6.0 array.

From a total of 927 women (333 from SFBCS and 549 from NC-BCFR), we excluded from all analyses 28 (3%) cases with missing data on patient characteristics.

The study was approved by the Institutional Review Board for Human Subjects at all institutions. All participants signed a written informed consent.

Data collection

Breast cancer cases for both studies were identified through the Greater Bay Area Cancer Registry, which also provided information on tumor characteristics (stage, grade, size, estrogen and progesterone receptor status) and treatment during the first four months after diagnosis. Through linkage with the cancer registry, vital status was obtained as of October 31, 2011. Breast cancer-specific deaths were derived from the underlying cause of death on the death certificate based on ICD-9 (9174-175) or ICD-10 (C50) codes.

The two studies collected information on demographic and cultural background (e.g., place of birth, age at migration, language use), body size, and reproductive history, using similar structured questionnaires that were administered by trained bilingual and bicultural interviewers in Spanish or English (18, 19). Given that the study questionnaires did not collect information on individual level socioeconomic status-related variables, except education, neighborhood level socioeconomic status was estimated using the Yost census-based principal component analysis approach (21). Briefly, the Yost composite index is obtained by geocoding the individual's residential address at diagnosis to a census block and assigning each individual a previously developed block-specific socioeconomic status index value. This index is based on a principal component analysis of census data on education,

income and occupation. The first principal component represents a single composite measure of socioeconomic status. This measure of socioeconomic status has been previously associated with breast cancer incidence and risk and has been reported to complement the effect of individual-level education (21–23).

We used available information on place of origin of self, parents and grandparents to create four region categories: Central America (Mexico, El Salvador, Guatemala, Honduras, Panama and Costa Rica), Caribbean (Puerto Rico, Cuba and Dominican Republic), South America (Venezuela, Colombia, Peru, Bolivia, Brazil, Argentina, Chile, Ecuador) and U.S. (which includes all individuals that self-identified as Latinas/Hispanic but responded “U.S.” or other category to the question about origin of self and of all parents and grandparents).

Estimation of individual genetic ancestry

We estimated global individual ancestry as the average locus-specific ancestry across 59,211 loci for each individual. Locus-specific ancestry estimates obtained with the HAPMIX software (24) were available from a previous genome-wide genotyping effort described elsewhere (13).

Statistical Analysis

We used one-way analyses of variance or t tests to explore the associations between genetic ancestry and patient characteristics, tumor characteristics and treatment received. African ancestry did not follow a normal distribution and therefore was log transformed (natural logarithm) for statistical analyses.

We used multivariate Cox proportional hazards models to evaluate the association between Indigenous American ancestry and breast cancer-specific mortality. Follow-up time in years was calculated as the difference between date of last follow-up and date of diagnosis.

We first conducted an analysis with Indigenous American ancestry defined in the model as a continuous variable, which was scaled so that the coefficient from the Cox model would reflect the effect on survival for each 25% increase in Indigenous American ancestry. Based on prior knowledge about possible prognostic factors, the analyses were adjusted for neighborhood socioeconomic status (continuous), education (<8 years, 8–11 years, high school graduate, and some college or higher), body mass index (<30, 30 kg/m²), family history of breast cancer in first-degree relatives (yes, no), place of birth [foreign-born (yes, no) and region of origin (Central America, Caribbean, South America, U.S.)], African ancestry (continuous), and first language spoken (English, English and another language, no English). We included tumor characteristics to test if the association between genetic ancestry and mortality could be due to the association between genetic ancestry and tumor stage (localized, regional extension or nodes only, regional extension and nodes, remote extension), grade (highly differentiated, moderately differentiated, poorly differentiated or Undifferentiated), size (<20mm, >20mm) and hormone receptor status (ER+/PR+, ER+/PR–, ER–/PR+ and ER–/PR). We also tested if the observed association could be due to differences in treatment received during the first four months after diagnosis [type of surgery (none, lumpectomy or mastectomy); radiation (yes, no), and chemotherapy (yes, no)].

To further explore the association between genetic ancestry and breast cancer-specific survival we created a categorical ancestry variable with four categories (0–25%, 25–50%, 50–75% and 75–100% Indigenous American ancestry) and used Kaplan-Meier survival estimates to visualize the ancestry effect. Based on the results of the Kaplan-Meier survival curves analysis, we conducted the Cox model analyses with ancestry defined as a dichotomous variable (<50% vs. ≥50% Indigenous American ancestry).

The program STATA 12.1 was used for analyses and all tests were two-sided (25).

Multiple Imputation

We imputed values for missing tumor stage, grade and hormone receptor status data in order to maintain the sample size when running the fully adjusted model. We used multiple imputation as theoretically described by Rubin (26) and implemented in the STATA 12.1 software (25). Imputation is a simple Monte Carlo technique that replaces missing values with imputed values and combines statistics from analysis conducted with multiple sets of imputed data. This approach is appropriate if the missing values are missing at random, that is, the missing data values carry no information about probabilities of missingness. Data on tumor characteristics are systematically collected by the Greater Bay Area Cancer Registry and therefore “missing at random” is a reasonable assumption. We specified 50 imputed datasets to achieve appropriate levels of variation in our estimates and 1,000 iterations (10 burn-in).

Results

Association between genetic ancestry and potential predictors for breast cancer-specific mortality

Indigenous American genetic ancestry was significantly higher among women diagnosed at age ≥50 years, those who had a lower level of education, did not speak English as their first language, were born outside the U.S., and had a low socioeconomic status (Table 1). Indigenous American ancestry was lower among women with family history of breast cancer and varied by region of origin, ranging from 16% among women of Caribbean origin to 42% among women from Central America. We did not observe significant differences in genetic ancestry between obese and non-obese women, women with different tumor characteristics (including the “missing” category), or by treatment received. We also investigated the association between African genetic ancestry and these variables. African ancestry was slightly lower among women with higher education and higher socioeconomic status, those born in the U.S., and those who spoke English as their first language. African ancestry was significantly higher among women of Caribbean origin (median of 12%) and lowest among women from South America (median of 5%).

Association between genetic ancestry and breast cancer-specific mortality

Indigenous American ancestry was statistically significantly associated with breast cancer-specific mortality in an analysis only adjusted by study [hazard ratio (HR) per every 25% increase in Indigenous American ancestry = 1.57, 95%CI 1.08–2.29] (Table 2). In bivariate analyses of other possible predictors of breast cancer specific-mortality we observed a

higher hazard of mortality among women who were diagnosed at a younger age, had a higher tumor stage, higher grade, larger tumor size, hormone receptor negative tumors, had a mastectomy or no surgery instead of a lumpectomy, and those who received chemotherapy (Table 2). In the fully adjusted model only Indigenous American ancestry (HR per 25% increase in Indigenous American ancestry: 1.75, 95%CI 1.12–2.74), region of origin, tumor stage, tumor size, and type of surgery were statistically significantly associated with mortality (Table 3). Caribbean origin was associated with a statistically significant increased risk of mortality when compared to women of Central American origin (HR 3.62, 95%CI 1.08–12.09) only in the fully adjusted model. In an analysis stratified by study, the results showed no evidence of heterogeneity (for NC-BCFR: HR 1.77, 95%CI 0.96–3.25; for SFBGS: HR 1.63, 95% CI 0.81–3.27, study-ancestry interaction p value 0.915).

We also ran the fully adjusted Cox model in a subsample of 545 individuals without missing data (47 breast cancer-specific deaths) to assess if the observed statistically significant association between Indigenous American ancestry and survival could be due to improper adjustment as a result of the imputation. The results of the analysis with a reduced set of samples showed no indication that the association between ancestry and survival might be due to improper adjustment when using imputed data (Unadjusted HR: 1.43, 95%CI 0.88–2.33. Adjusted HR: 1.54, 95%CI 0.91–2.63).

We further explored the relationship between genetic ancestry and survival by categorizing genetic ancestry and conducting Kaplan-Meier survival estimates. We first conducted an analysis that included four ancestry categories: 0–25%, 25–50%, 50–75% and 75–100% to be consistent with the magnitude of the unit change defined for the continuous ancestry model. This analysis strongly suggested that there was no difference in survival between individuals with 0–25% and 25–50% Indigenous American ancestry but that survival differed between individuals within these two lower ancestry groups and the individuals with 50–75% and 75–100% Indigenous American ancestry (Log-rank test P value = 0.015) (Figure 1.A). To obtain a HR estimate that corresponded more closely with the observed results, we dichotomized genetic ancestry (<50% vs. >50% of Indigenous American ancestry) and found a HR of 1.89 (95%CI 1.10–3.24) (Log-rank test P value=0.003) (Table 3 & Figure 1.B).

DISCUSSION

We found that among U.S. Latinas from Northern California, Indigenous American ancestry was significantly associated with breast cancer-specific mortality. The association remained statistically significant after adjustment for age at diagnosis, body mass index, family history of breast cancer, place of birth, African ancestry, language use, socioeconomic status, education, tumor grade, stage, size, hormone receptor status, and treatment received in the first four months after diagnosis. The adjusted hazard ratio for every 25% increase in Indigenous American ancestry was 1.75 (95%CI: 1.12–2.74). Kaplan-Meier survival estimates based on a categorized ancestry variable strongly suggested that the effect of ancestry was non-monotonic and that the difference in survival was most pronounced when comparing women with >50% vs. <50% Indigenous American ancestry. The results showed that the hazard of mortality was twice as large among women with approximately half their

genomes being of Indigenous American origin compared to women with less Indigenous American ancestry.

The association between genetic ancestry and survival in our study was in the expected direction based on the previously described increased risk of breast cancer mortality among U.S. Latinas compared to non-Latina Whites (3, 27, 28). The association persisted after adjusting for multiple prognostic factors, which is compatible with the hypothesis of an underlying genetic component linking genetic ancestry and mortality in Latinas. However, it is also possible that our model did not include all relevant prognostic factors, thus, a role for non-genetic factors cannot be discarded. Non-Latina Whites tend to have higher level of screening and adequate treatment compared to other racial/ethnic groups (29, 30). In previous studies, the disparity in outcomes between non-Latina Whites and other groups, including Latinas, have been explained by differences in access to care, socioeconomic status and level of education (29–32). In the present study, the disparity in outcome by genetic ancestry was still observed after controlling for socioeconomic status, education, language use, place of birth, and treatment received. However, access to care and adherence to treatment beyond the first four months after diagnosis might not be properly captured by these available variables. Hormone therapy is highly effective for women with ER-positive disease, but inadequate adherence may substantially reduce its efficacy (33, 34). Furthermore, the type of hormone therapy may impact outcome. For example, a randomized clinical trial showed that more patients who received neo-adjuvant therapy with aromatase inhibitors had complete or partial responses than patients who received tamoxifen (35). Due to lower socioeconomic status, education and other related factors, women with higher Indigenous American ancestry may have been prescribed aromatase inhibitors less frequently. Future studies should explore the possibility that these factors might mediate the relationship between genetic ancestry and breast cancer-specific mortality among U.S. Latinas.

An interesting observation is that even though region of origin was not associated with mortality in a bivariate model, Caribbean origin was statistically significantly associated with increased risk of mortality in a fully adjusted model. Reports using SEER data also found increased mortality among Latina breast cancer cases of Caribbean origin (36). This association should be explored in future studies.

Our study did not find a significant association between African ancestry and breast cancer-specific mortality. African American women in the U.S. have significantly higher breast cancer mortality than any other group. We have previously reported that among African American women, those with higher European ancestry are more likely to be diagnosed with ER-positive breast cancer, which tends to have a better prognosis. However, the proportion of African ancestry in our sample was low (~8%), which decreased our power to confirm such an association.

Family studies among women of European descent have suggested that breast cancer prognosis might have an inherited component (16, 17). Several studies have reported associations between SNPs within candidate genes and breast cancer survival (37–40), but most variants investigated are of unknown functional significance, and few have been

confirmed (41, 42). Among genetic variants known to be associated with breast cancer susceptibility, only one locus in the *TOX3* gene was associated with survival at a level of statistical significance ($p < 0.01$) (41). These results suggest that breast cancer survival may be influenced by a set of genetic variants different from those influencing cancer risk susceptibility. Few genome-wide association studies have been conducted to date to identify genotypes associated with clinical outcomes such as cancer recurrence, or overall survival (43–45). Shu et al. (43) reported a variant in the *RAD51L1* gene and one in a locus on chromosome 16 to be associated with all-cause mortality among breast cancer patients from Shanghai. They replicated their results among European women from the Nurses' Health Study. One of the studies conducted among European women reported a SNP in the *OCA2* gene associated with all-cause mortality among women with estrogen receptor-negative tumors (44), but a second study conducted as part of the Cancer Genetic Markers of Susceptibility (CGEMS) Study, found no SNPs with a statistically significant association with breast cancer survival (46). A two-stage study among women with early-onset breast cancer in the United Kingdom reported a SNP associated with breast cancer-specific survival upstream of the *ARRDC3* locus (45). The association between genetic ancestry and breast cancer-specific mortality among U.S. Latinas, after adjusting for multiple potential confounders, raises the possibility that genetic variants that affect survival might differ between populations. Recent results in other cancers, such as neuroblastoma (47) or acute lymphoblastic leukemia (48), support this possibility.

Our findings may also be due to differences in the subtype of breast cancer that occurs among women with higher Indigenous American ancestry vs. women with higher European ancestry. Although we did not detect any association between ER/PR status and genetic ancestry, there is substantial heterogeneity among ER-positive breast cancers that can be identified by a variety of gene expression and other molecular markers that is also associated with prognosis (49). It is possible that women with higher Indigenous American ancestry are at higher risk for one of the more aggressive ER-positive subset of tumors, or that women with higher European ancestry are at increased risk for a less aggressive ER-positive subset of tumors. We have previously demonstrated that Latina women with higher European ancestry are at higher risk for overall breast cancer. If women with higher European ancestry are at increased overall risk for breast cancer, but they are at particularly higher risk for a less aggressive subtype, this could account for both our current results and our previously reported results.

This is the first study looking at the relationship between genetic ancestry and breast cancer-specific survival among U.S. Latinas. Replication of the observed association in an independent sample of Latinas should be of high priority. Given the relatively high survival of women who develop breast cancer, it has been a challenge to find a cohort of Latina cases with enough events to test the presently observed association. We had access to genetic ancestry estimates, ER/PR status, age at diagnosis and survival information for 335 women with breast cancer from the Multiethnic Cohort study (MEC). There were a total of 35 deaths from breast cancer and an average of 6 years of follow up time. Even though the sample was small and we did not have information on many of the potential confounders, results were consistent with our findings (Unadjusted HR $< 50\%$ vs. $> 50\%$ Indigenous

American ancestry=1.52, 95%CI 0.69–3.35; HR adjusted for age at diagnosis and ER/PR status=1.45, 95%CI 0.65–3.24).

We observed that the probability of survival was slightly higher in the SFBCS than in the MEC. This difference is likely to be due to the smaller proportion of advanced stage disease patients represented in the study (~5%) compared to the expected proportion based on the general patient population (7–10%). However, this potential limitation of the study is unlikely to have affected the association between genetic ancestry and survival because our analyses showed that the association was independent of tumor stage, grade, tumor size or hormone status. When we limited the analysis to women with early stage disease, the HR for the <50% vs. 50% Indigenous American ancestry model was still 2.09 (P value 0.011), which was very similar to the HR of 2.05 for the complete data set.

In the U.S., Latinos are the largest and youngest minority group, accounting for 15% of the Nation's total population. Currently, there are ~42 million Latinos in the U.S., not including residents of Puerto Rico. U.S. Latinas, despite their relatively low incidence of breast cancer, tend to be diagnosed at a young age, with aggressive tumors and are more likely to die from the disease than non-Latina White women. Our study explored the difference in breast cancer-specific survival among Latinas with different proportions of Indigenous, European and African ancestry. We showed that Latina women with 50% Indigenous American ancestry have twice the hazard of breast cancer-specific mortality than Latina women with <50% Indigenous American ancestry. This important disparity could be due to non-genetic factors that were unaccounted for in the present study or to genetic differences. Identifying novel genetic variants that are associated with survival in this population might contribute to the understanding of the biological mechanisms that lead to worse breast cancer prognosis in this group as well as to the development of new treatments. Identification of putative modifiable socioeconomic or cultural factors that affect breast cancer-specific survival among Latinas with high Indigenous American ancestry could lead to the elimination of the observed disparity.

Acknowledgments

Funding

Work for this study was supported by the National Cancer Institute [CA 160607 to L.F. and CA120120 to E.Z.], the Center for Aging in Diverse Communities (CADC) under the Resource Centers for Minority Aging Research program by the National Institute on Aging [P30-AG15272] and the National Cancer Institute grant from the Special Population Network program to University of Texas, San Antonio [Redes En Acción # U01CA86117]. The San Francisco Bay Area Breast Cancer Study was supported by the National Cancer Institute [CA63446 and CA77305], the U.S. Department of Defense [DAMD17-96-1-6071] and the California Breast Cancer Research Program [7PB-0068]. The Northern California Breast Cancer Family Registry was supported by grant UM1 CA164920 from the National Cancer Institute. The content of this manuscript does not necessarily reflect the views or policies of the National Cancer Institute or any of the collaborating centers in the Breast Cancer Family Registry (BCFR), nor does mention of trade names, commercial products, or organizations imply endorsement by the US Government or the BCFR. The collection of cancer incidence data used in this study was supported by the California Department of Public Health as part of the statewide cancer reporting program mandated by California Health and Safety Code Section 103885; the National Cancer Institute's Surveillance, Epidemiology and End Results Program under contract HHSN261201000140C awarded to the Cancer Prevention Institute of California, contract HHSN261201000035C awarded to the University of Southern California, and contract HHSN261201000034C awarded to the Public Health Institute; and the Centers for Disease Control and Prevention's National Program of Cancer Registries, under agreement U58DP003862-01 awarded to the California Department of Public Health. The ideas and opinions expressed herein are those of the author(s) and endorsement

by the State of California, Department of Public Health the National Cancer Institute, and the Centers for Disease Control and Prevention or their Contractors and Subcontractors is not intended nor should be inferred.

The authors would like to thank the study participants.

References

1. Smigal C, Jemal A, Ward E, Cokkinides V, Smith R, Howe HL, et al. Trends in breast cancer by race and ethnicity: update 2006. *CA Cancer J Clin.* 2006; 56:168–183. [PubMed: 16737949]
2. Jemal A, Ward E, Thun MJ. Recent trends in breast cancer incidence rates by age and tumor characteristics among U.S. women. *Breast Cancer Res.* 2007; 9:R28. [PubMed: 17477859]
3. Li CI, Malone KE, Daling JR. Differences in breast cancer stage, treatment, and survival by race and ethnicity. *Arch Intern Med.* 2003; 163:49–56. [PubMed: 12523916]
4. Newman LA, Mason J, Cote D, Vin Y, Carolin K, Bouwman D, et al. African-American ethnicity, socioeconomic status, and breast cancer survival: a meta-analysis of 14 studies involving over 10,000 African-American and 40,000 White American patients with carcinoma of the breast. *Cancer.* 2002; 94:2844–2854. [PubMed: 12115371]
5. Sprague BL, Trentham-Dietz A, Gangnon RE, Ramchandani R, Hampton JM, Robert SA, et al. Socioeconomic status and survival after an invasive breast cancer diagnosis. *Cancer.* 2011; 117:1542–1551. [PubMed: 21425155]
6. Newman LA, Griffith KA, Jatoi I, Simon MS, Crowe JP, Colditz GA. Meta-analysis of survival in African American and white American patients with breast cancer: ethnicity compared with socioeconomic status. *J Clin Oncol.* 2006; 24:1342–1349. [PubMed: 16549828]
7. Bradley CJ, Given CW, Roberts C. Race, socioeconomic status, and breast cancer treatment and survival. *J Natl Cancer Inst.* 2002; 94:490–496. [PubMed: 11929949]
8. Boyer-Chamhard A, Taylor TH, Anton-Culver H. Survival differences in breast cancer among racial/ethnic groups: a population-based study. *Cancer Detect Prev.* 1999; 23:463–473. [PubMed: 10571656]
9. Via M, Gignoux CR, Roth LA, Fejerman L, Galanter J, Choudhry S, et al. History shaped the geographic distribution of genomic admixture on the island of Puerto Rico. *PLoS One.* 2011; 6:e16513. [PubMed: 21304981]
10. Silva-Zolezzi I, Hidalgo-Miranda A, Estrada-Gil J, Fernandez-Lopez JC, Uribe-Figueroa L, Contreras A, et al. Analysis of genomic diversity in Mexican Mestizo populations to develop genomic medicine in Mexico. *Proc Natl Acad Sci U S A.* 2009; 106:8611–8616. [PubMed: 19433783]
11. Parra EJ, Kittles RA, Shriver MD. Implications of correlations between skin color and genetic ancestry for biomedical research. *Nat Genet.* 2004; 36:S54–S60. [PubMed: 15508005]
12. Kosoy R, Nassir R, Tian C, White PA, Butler LM, Silva G, et al. Ancestry informative marker sets for determining continental origin and admixture proportions in common populations in America. *Hum Mutat.* 2009; 30:69–78. [PubMed: 18683858]
13. Fejerman L, Chen GK, Eng C, Huntsman S, Hu D, Williams A, et al. Admixture mapping identifies a locus on 6q25 associated with breast cancer risk in US Latinas. *Hum Mol Genet.* 2012
14. Fejerman L, John EM, Huntsman S, Beckman K, Choudhry S, Perez-Stable E, et al. Genetic ancestry and risk of breast cancer among U.S. Latinas. *Cancer Res.* 2008; 68:9723–9728. [PubMed: 19047150]
15. Fejerman L, Romieu I, John EM, Lazcano-Ponce E, Huntsman S, Beckman KB, et al. European ancestry is positively associated with breast cancer risk in Mexican women. *Cancer Epidemiol Biomarkers Prev.* 2010; 19:1074–1082. [PubMed: 20332279]
16. Hartman M, Lindstrom L, Dickman PW, Adami HO, Hall P, Czene K. Is breast cancer prognosis inherited? *Breast Cancer Res.* 2007; 9:R39. [PubMed: 17598882]
17. Lindstrom LS, Hall P, Hartman M, Wiklund F, Gronberg H, Czene K. Familial concordance in cancer survival: a Swedish population-based study. *Lancet Oncol.* 2007; 8:1001–1006. [PubMed: 17921068]

18. John EM, Horn-Ross PL, Koo J. Lifetime physical activity and breast cancer risk in a multiethnic population: the San Francisco Bay area breast cancer study. *Cancer Epidemiol Biomarkers Prev.* 2003; 12:1143–1152. [PubMed: 14652273]
19. John EM, Phipps AI, Davis A, Koo J. Migration history, acculturation, and breast cancer risk in Hispanic women. *Cancer Epidemiol Biomarkers Prev.* 2005; 14:2905–2913. [PubMed: 16365008]
20. John EM, Hopper JL, Beck JC, Knight JA, Neuhausen SL, Senie RT, et al. The Breast Cancer Family Registry: an infrastructure for cooperative multinational, interdisciplinary and translational studies of the genetic epidemiology of breast cancer. *Breast Cancer Res.* 2004; 6:R375–R389. [PubMed: 15217505]
21. Yost K, Perkins C, Cohen R, Morris C, Wright W. Socioeconomic status and breast cancer incidence in California for different race/ethnic groups. *Cancer Causes Control.* 2001; 12:703–711. [PubMed: 11562110]
22. Keegan TH, John EM, Fish KM, Alfaro-Velcamp T, Clarke CA, Gomez SL. Breast cancer incidence patterns among California Hispanic women: differences by nativity and residence in an enclave. *Cancer Epidemiol Biomarkers Prev.* 2010; 19:1208–1218. [PubMed: 20447917]
23. Webster TF, Hoffman K, Weinberg J, Vieira V, Aschengrau A. Community- and individual-level socioeconomic status and breast cancer risk: multilevel modeling on Cape Cod, Massachusetts. *Environ Health Perspect.* 2008; 116:1125–1129. [PubMed: 18709175]
24. Price AL, Tandon A, Patterson N, Barnes KC, Rafaels N, Ruczinski I, et al. Sensitive detection of chromosomal segments of distinct ancestry in admixed populations. *PLoS Genet.* 2009; 5:e1000519. [PubMed: 19543370]
25. StataCorp. Stata Statistical Software: Release 12.1. College Station, TX: StataCorp LP; 2011.
26. Rubin DB, Schenker N. Multiple imputation in health-care databases: an overview and some applications. *Stat Med.* 1991; 10:585–598. [PubMed: 2057657]
27. Campbell JB. Breast cancer-race, ethnicity, and survival: a literature review. *Breast Cancer Res Treat.* 2002; 74:187–192. [PubMed: 12186379]
28. Le Marchand L. Ethnic variation in breast cancer survival: a review. *Breast Cancer Res Treat.* 1991; 18(Suppl 1):S119–S126. [PubMed: 1873548]
29. Austin LT, Ahmad F, McNally MJ, Stewart DE. Breast and cervical cancer screening in Hispanic women: a literature review using the health belief model. *Womens Health Issues.* 2002; 12:122–128. [PubMed: 12015184]
30. Freedman RA, Virgo KS, He Y, Pavluck AL, Winer EP, Ward EM, et al. The association of race/ethnicity, insurance status, and socioeconomic factors with breast cancer care. *Cancer.* 2011; 117:180–189. [PubMed: 20939011]
31. Bickell NA, Wang JJ, Oluwole S, Schrag D, Godfrey H, Hiotis K, et al. Missed opportunities: racial disparities in adjuvant breast cancer treatment. *J Clin Oncol.* 2006; 24:1357–1362. [PubMed: 16549830]
32. Coughlin SS, Uhler RJ. Breast and cervical cancer screening practices among Hispanic women in the United States and Puerto Rico, 1998–1999. *Prev Med.* 2002; 34:242–251. [PubMed: 11817921]
33. Hershman DL, Shao T, Kushi LH, Buono D, Tsai WY, Fehrenbacher L, et al. Early discontinuation and non-adherence to adjuvant hormonal therapy are associated with increased mortality in women with breast cancer. *Breast Cancer Res Treat.* 2011; 126:529–537. [PubMed: 20803066]
34. McCowan C, Shearer J, Donnan PT, Dewar JA, Crilly M, Thompson AM, et al. Cohort study examining tamoxifen adherence and its relationship to mortality in women with breast cancer. *Br J Cancer.* 2008; 99:1763–1768. [PubMed: 18985046]
35. Masuda N, Sagara Y, Kinoshita T, Iwata H, Nakamura S, Yanagita Y, et al. Neoadjuvant anastrozole versus tamoxifen in patients receiving goserelin for premenopausal breast cancer (STAGE): a double-blind, randomised phase 3 trial. *Lancet Oncol.* 2012; 13:345–352. [PubMed: 22265697]
36. Chakraborty S, Smith L, Ganti AK, Bonthu N, Shimizu T, Batra SK. Breast cancer survival of Hispanic women in the USA is influenced by country of origin. *Asia Pac J Clin Oncol.* 2013

37. Long JR, Kataoka N, Shu XO, Wen W, Gao YT, Cai Q, et al. Genetic polymorphisms of the CYP19A1 gene and breast cancer survival. *Cancer Epidemiol Biomarkers Prev.* 2006; 15:2115–2122. [PubMed: 17119036]
38. Lu H, Shu XO, Cui Y, Kataoka N, Wen W, Cai Q, et al. Association of genetic polymorphisms in the VEGF gene with breast cancer survival. *Cancer Res.* 2005; 65:5015–5019. [PubMed: 15958542]
39. Deming SL, Ren Z, Wen W, Shu XO, Cai Q, Gao YT, et al. Genetic variation in IGF1, IGF-1R, IGFALS, and IGFBP3 in breast cancer survival among Chinese women: a report from the Shanghai Breast Cancer Study. *Breast Cancer Res Treat.* 2007; 104:309–319. [PubMed: 17063263]
40. Zhang X, Shu XO, Cai Q, Ruan Z, Gao YT, Zheng W. Functional plasminogen activator inhibitor-1 gene variants and breast cancer survival. *Clin Cancer Res.* 2006; 12:6037–6042. [PubMed: 17062678]
41. Fasching PA, Pharoah PD, Cox A, Nevanlinna H, Bojesen SE, Karn T, et al. The role of genetic breast cancer susceptibility variants as prognostic factors. *Hum Mol Genet.* 2012
42. Goode EL, Dunning AM, Kuschel B, Healey CS, Day NE, Ponder BA, et al. Effect of germ-line genetic variation on breast cancer survival in a population-based study. *Cancer Res.* 2002; 62:3052–3057. [PubMed: 12036913]
43. Shu XO, Long J, Lu W, Li C, Chen WY, Delahanty R, et al. Novel Genetic Markers of Breast Cancer Survival Identified by a Genome-Wide Association Study. *Cancer Res.* 2012; 72:1182–1189. [PubMed: 22232737]
44. Azzato EM, Tyrer J, Fasching PA, Beckmann MW, Ekici AB, Schulz-Wendtland R, et al. Association between a germline OCA2 polymorphism at chromosome 15q13.1 and estrogen receptor-negative breast cancer survival. *J Natl Cancer Inst.* 2010; 102:650–662. [PubMed: 20308648]
45. Rafiq S, Tapper W, Collins A, Khan S, Politopoulos I, Gerty S, et al. Identification of Inherited Genetic Variations Influencing Prognosis in Early-Onset Breast Cancer. *Cancer Res.* 2013; 73:1883–1891. [PubMed: 23319801]
46. Azzato EM, Pharoah PD, Harrington P, Easton DF, Greenberg D, Caporaso NE, et al. A genome-wide association study of prognosis in breast cancer. *Cancer Epidemiol Biomarkers Prev.* 2010; 19:1140–1143. [PubMed: 20332263]
47. Gamazon ER, Pinto N, Konkashbaev A, Im HK, Diskin SJ, London WB, et al. Trans-population analysis of genetic mechanisms of ethnic disparities in neuroblastoma survival. *J Natl Cancer Inst.* 2013; 105:302–309. [PubMed: 23243203]
48. Yang JJ, Cheng C, Devidas M, Cao X, Fan Y, Campana D, et al. Ancestry and pharmacogenomics of relapse in acute lymphoblastic leukemia. *Nat Genet.* 2011; 43:237–241. [PubMed: 21297632]
49. O'Brien KM, Cole SR, Tse CK, Perou CM, Carey LA, Foulkes WD, et al. Intrinsic breast tumor subtypes, race, and long-term survival in the Carolina Breast Cancer Study. *Clin Cancer Res.* 2010; 16:6100–6110. [PubMed: 21169259]

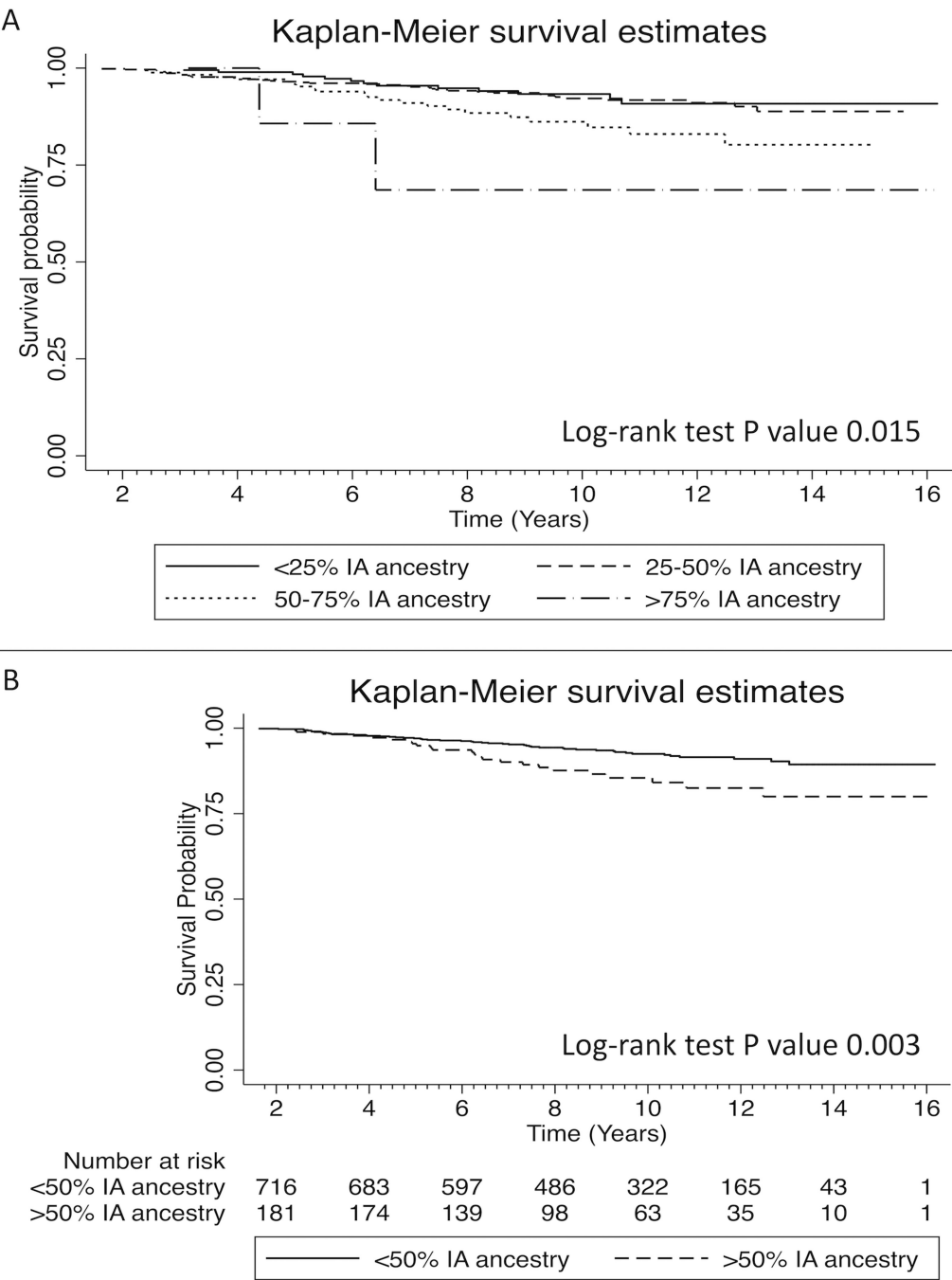


Figure 1.
A) Kaplan-Meier survival plot comparing 899 U.S. Latina breast cancer patients in 4 categories of Indigenous American (IA) ancestry (<25%, 25–50, 50–75, 75–100). The X and Y axes represent follow-up time in years and the percentage of survival, respectively. The associated log-rank p value is 0.015. B) Kaplan-Meier survival plot comparing 899 U.S. Latina breast cancer patients in 2 categories of Indigenous American (IA) ancestry (<50% & >50%). The associated log-rank p value is 0.003.

Median (Interquartile range) percent genetic ancestry by patient and tumor characteristics and treatment among U.S. Latina breast cancer cases ^a

Table 1

Patient Characteristics	Categories	N (%)	Ind. Am. Anc.	P value	Afr. anc. ^c	P value
Age at diagnosis	50 years	458 (51)	41 (19)	<0.001	7 (3)	0.354
	>50 years	441 (49)	38 (20)		7 (3)	
Obesity (BMI ≥ 30 kg/m ²)	No	587 (65)	39 (20)	0.534	7 (3)	0.415
	Yes	312 (35)	40 (18)		7 (3)	
Family history of breast cancer	No	752 (84)	40 (19)	<0.001	7 (3)	0.437
	Yes	147 (16)	35 (21)		7 (3)	
Education	< 8 years	164 (18)	43 (17)	<0.001	7 (4)	0.003
	8–11 years	167 (19)	43 (17)		7 (3)	
	High school graduate	250 (28)	40 (16)		7 (3)	
	Some college or higher	318 (35)	34 (25)		6 (3)	
Language first spoken	English	252 (28)	31 (25)	<0.001	6 (3)	0.015
	English & other language	60 (7)	38 (22)		6 (3)	
	Other language	587 (65)	42 (16)		7 (3)	
Foreign-born	Yes	453 (50)	42 (18)	<0.001	7 (4)	<0.001
	No	446 (50)	36 (22)		6 (3)	
Region of origin	Central America	663 (74)	42 (16)	<0.001	7 (3)	<0.001
	Caribbean	48 (5)	16 (8)		12 (11)	
	South America	52 (6)	41 (33)		5 (3)	
	U.S.	136 (15)	29 (28)		6 (2)	
Socioeconomic status (SES) ^b	1 st Quintile (lowest SES)	63 (7)	44 (13)	<0.001	7 (3)	0.021
	2 nd Quintile	135 (15)	40 (17)		7 (2)	
	3 rd Quintile	201 (22)	41 (20)		7 (4)	
	4 th Quintile	248 (28)	38 (19)		7 (3)	
	5 th Quintile (Highest SES)	252 (28)	35 (21)		6 (3)	

Patient Characteristics	Categories	N (%)	Ind. Am. Anc.	P value	Afr. anc. ^c	P value
Clinical Characteristics						
Tumor stage	Localized	490 (55)	38 (19)	0.807	7 (3)	0.239
	Regional extension or nodes	235 (26)	39 (19)		7 (3)	
	Regional extension & nodes	22 (2)	35 (21)		6 (3)	
	Remote extension	15 (2)	36 (23)		6 (3)	
	Missing	137 (15)	42 (20)		7 (3)	
Tumor grade	Highly differentiated	141 (16)	42 (18)	0.968	7 (3)	0.335
	Moderately differentiated	317 (35)	38 (19)		7 (3)	
	Poorly differentiated & Undifferentiated	349 (39)	39 (18)		7 (3)	
	Missing	92 (10)	38 (25)		7 (4)	
Tumor ER/PR status	ER+/PR+	389 (43)	39 (20)	0.535	7 (3)	0.224
	ER+/PR-	87 (10)	38 (16)		7 (3)	
	ER-/PR+	13 (2)	38 (11)		6 (2)	
	ER-/PR-	156 (17)	40 (18)		7 (3)	
	Missing	254 (28)	39 (22)		7 (3)	
Tumor Size	20mm	432	38 (20)	0.069	7 (4)	0.475
	>20mm	268	41 (19)		7 (3)	
	Missing	199	39 (18)		7 (3)	
Surgery	No surgery	14 (2)	32 (23)	0.204	6 (2)	0.624
	Lumpectomy	519 (58)	39 (19)		7 (3)	
	Mastectomy	366 (40)	40 (18)		7 (3)	
Radiation	No	357 (40)	39 (19)	0.311	7 (3)	0.680
	Yes	542 (60)	39 (19)		7 (3)	
Chemotherapy	No	351 (39)	38 (20)	0.083	7 (4)	0.201
	Yes	542 (60)	40 (19)		7 (3)	
	Missing	6 (1)	25 (22)		7 (3)	

Afr. anc: African ancestry; Ind. Am. anc.: Indigenous American ancestry; ER: Estrogen Receptor; PR: Progesterone Receptor

Author Manuscript

Author Manuscript

Author Manuscript

Author Manuscript

^dBased on 899 breast cancer cases

^qQuintiles are based on the distribution of socioeconomic status index values among unaffected controls from the SFBCS. The unbalanced proportion of cases within each quintile in the present analysis reflects the observation that breast cancer cases tend to have higher socioeconomic status than breast cancer controls.

^cFor test of ANOVA African ancestry was log transformed to approximate normality

Table 2

Bivariate Cox proportion hazards models evaluating the association between breast cancer-specific mortality and patient and tumor characteristics and treatment among U.S. Latina breast cancer cases ^a

Patient Characteristics	Categories	Hazard Ratio	95% CI	P value ^b
Indigenous American ancestry (25% unit)		1.57	1.08 – 2.29	0.019
African ancestry (25% unit)		0.57	0.15 – 2.03	0.375
Age at diagnosis (5 yrs)		0.85	0.76 – 0.94	0.003
Socioeconomic status (Yost Index)		0.86	0.66 – 1.12	0.262
Obese (≥ 30 kg/m ²)	No	1.0		
	Yes	1.28	0.80 – 2.04	0.301
Family history of breast cancer	No	1.0		
	Yes	0.52	0.24 – 1.15	0.108
Foreign-born	No	1.0		
	Yes	1.1	0.70 – 1.73	0.689
Region of origin	Central America	1.0		
	Caribbean	1.1	0.44–2.76	0.833
	South America	1.12	0.45–2.81	0.802
	U.S.	0.77	0.38–1.56	0.470
Education	<8 years	1.0		
	8 to 11 years	0.87	0.43 – 1.74	0.688
	High school graduate	0.71	0.37 – 1.36	0.298
	Some college or higher	0.69	0.37 – 1.28	0.236
Language first spoken	Other language	1.0		
	English & other language	1.04	0.44 – 2.46	0.931
	English	1.28	0.78 – 2.10	0.334
Clinical Characteristics				
Stage, N=762	Localized	1.0		
	Regional extension or nodes	3.55	2.07 – 6.09	<0.001
	Regional extension & nodes	7.09	2.86 – 17.63	<0.001
	Remote extension	19.42	8.49 – 44.42	<0.001
Grade, N=807	Highly differentiated	1.0		
	Moderately differentiated	1.7	0.69 – 4.20	0.251
	Poorly differentiated & Undifferentiated	2.83	1.19 – 6.74	0.019
Hormone receptor status, N=645	ER+/PR+	1.0		
	ER+/PR–	2.59	1.38 – 4.88	0.003

Patient Characteristics	Categories	Hazard Ratio	95% CI	P value ^b
	ER-/PR+	2.12	0.50 – 8.93	0.305
	ER-/PR–	1.31	0.67 – 2.55	0.423
Tumor Size, N=700	20mm	1.0		
	>20mm	3.53	2.04–6.10	<0.001
Surgery	Lumpectomy	1.0		
	Mastectomy	2.57	1.58 – 4.16	<0.001
	No surgery	7.88	2.72 – 22.78	<0.001
Radiation	No	1.0		
	Yes	0.94	0.59 – 1.49	0.791
Chemotherapy, N=894	No	1.0		
	Yes	2.44	1.41 – 4.21	0.001

^aBased on 899 breast cancer cases and 75 deaths from breast cancer

^bAll models were adjusted by Study

Table 3

Cox proportion hazards evaluating the association between breast cancer-specific mortality and Indigenous American ancestry among U.S. Latina breast cancer cases ^a

Hazard ratios for Indigenous American ancestry	Categories	Hazard Ratio	95% CI	P value ^b
<i>Unadjusted</i>				
Indigenous American ancestry (25% unit)		1.57	1.08–2.29	0.019
Indigenous American ancestry (Categorical)	<50%	1.0		
	50%	2.05	1.26–3.34	0.004
<i>Adjusted</i>				
Indigenous American ancestry + A		1.58	1.03–2.42	0.035
Indigenous American ancestry + A + B		1.76	1.13–2.73	0.012
Indigenous American ancestry + A + B + C		1.75	1.12–2.74	0.014
Indigenous American ancestry + A	<50%	1.0		
	50%	1.88	1.12–3.17	0.017
Indigenous American ancestry + A + B	<50%	1.0		
	50%	1.83	1.07–3.11	0.027
Indigenous American ancestry + A + B + C	<50%	1.0		
	50%	1.89	1.10–3.24	0.021
Hazard ratios for covariates in the full model				
Personal characteristics = A				
Age at diagnosis (5 years)		0.95	0.83–1.08	0.413
Obesity (>30kg/m2)	No	1.0		
	Yes	1.29	0.76–2.19	0.337
Family history of breast cancer	No	1.0		
	Yes	0.74	0.32–1.72	0.484
Socioeconomic status (Yost Index)		0.89	0.65–1.23	0.487
Education	<8 years	1.0		
	8–11 years	0.73	0.32–1.67	0.459
	High school graduate	0.73	0.33–1.60	0.430
	Some college or higher	0.59	0.26–1.36	0.215
Language use	Other Language	1.0		
	English & other language	2.10	0.70–6.30	0.188
	English	2.37	0.99–5.64	0.052
Foreign-born	No	1.0		

Hazard ratios for Indigenous American ancestry	Categories	Hazard Ratio	95% CI	P value ^b
	Yes	1.17	0.51–2.67	0.717
African ancestry (25% unit)		0.40	0.07–2.29	0.301
Region of origin	Central America	1.0		
	Caribbean	3.62	1.08–12.09	0.037
	South America	1.22	0.45–3.30	0.695
	U.S.	0.66	0.29–1.50	0.327
Tumor characteristics = B				
Stage	Localized	1.0		
	Regional extension or nodes	2.69	1.39–5.18	0.003
	Regional extension & nodes	3.28	1.15–9.38	0.027
	Remote extension	17.02	5.84–49.58	<0.001
Grade	Highly differentiated	1.0		
	Moderately differentiated	1.12	0.43–2.91	0.816
	Poorly differentiated and Undifferentiated	1.41	0.52–3.82	0.503
Hormone receptor status	ER+/PR+	1.0		
	ER+/PR–	1.55	0.80–3.00	0.196
	ER–/PR+	1.54	0.34–7.05	0.579
	ER–/PR–	1.02	0.49–2.09	0.966
Tumor Size (mm)	20mm	1.0		
	>20mm	2.04	1.07–3.86	0.029
Treatment Information = C				
Surgery	Lumpectomy	1.0		
	Mastectomy	1.92	1.03–3.58	0.040
	No surgery	3.47	0.90–13.39	0.071
Radiation	No	1.0		
	Yes	1.24	0.68–2.27	0.483
Chemotherapy	No	1.0		
	Yes	0.83	0.40–1.70	0.604

ER: Estrogen Receptor; PR: Progesterone Receptor

^aBased on 899 breast cancer cases and 75 deaths from breast cancer

^bAll models were adjusted by study