ORIGINAL ARTICLE – BREAST ONCOLOGY

Disparities in Primary Breast Cancer Stage at Presentation Among Hispanic Subgroups

Nishwant Swami, MPH[1,2], Tiffany Nguyen, MD[3], Edward Christopher Dee, MD[4], Idalid Franco, MD, MPH[5], Yefri A. Baez, MD[6,7], Kaitlyn Lapen, MD[4], Lora Wang, MD[8], Neha Goel, MD[9], Brandon A. Mahal, MD[8], Oluwadamilola M. Fayanju, MD[10,11,12], Narjust Duma, MD[13], and Fumiko Chino, MD[4]

[1]University of Massachusetts Chan Medical School, Worcester, MA; [2]Harvard T.H. Chan School of Public Health, Boston, MA; [3]University of Miami Miller School of Medicine/Sylvester Comprehensive Cancer Center, Miami, FL; [4]Department of Radiation Oncology, Memorial Sloan Kettering Cancer Center, New York, NY; [5]Harvard Radiation Oncology Program, Harvard Medical School, Boston, MA; [6]Department of Urology, Brigham and Women's Hospital, Boston, MA; [7]Harvard Medical School, Boston, MA; [8]Department of Radiation Oncology, University of Miami Miller School of Medicine/ Sylvester Comprehensive Cancer Center, Miami, FL; [9]Department of Surgical Oncology, University of Miami Miller School of Medicine/Sylvester Comprehensive Cancer Center, Miami, FL; [10]Division of Endocrine and Oncologic Surgery, Department of Surgery, Perelman School of Medicine, The University of Pennsylvania, Philadelphia, PA; [11]Rena Rowan Breast Center, Abramson Cancer Center, Penn Medicine, Philadelphia, PA; [12]Penn Center for Cancer Care Innovation, The University of Pennsylvania, Philadelphia, PA; [13]Department of Medical Oncology, Dana-Farber Cancer Institute, Boston, MA

ABSTRACT

Background. Although the United States (US) Hispanic population consists of diverse communities, prior breast cancer studies often analyze this group in aggregate. Our aim was to identify differences in breast cancer stage at presentation in the US population, with a particular focus on Hispanic subgroups.

Methods. Data from the National Cancer Database (NCDB) from 2004 to 2017 were used to select women with primary breast cancer; individuals were disaggregated by racial and ethnic subgroup and Hispanic country of origin. Ordinal logistic regression was used to create adjusted odds ratios (aORs) with 95% confidence intervals (CIs), with higher odds representing presentation at later-stage breast cancer. Subgroup analysis was conducted based on tumor receptor status.

Results. Overall, among 2,282,691 women (5.2% Hispanic), Hispanic women were more likely to live in low-income and low-educational attainment neighborhoods, and were also more likely to be uninsured. Hispanic women were also more likely to present at later-stage primary breast cancer when compared with non-Hispanic White women (aOR 1.19, 95% CI 1.18–1.21; $p < 0.01$). Stage disparities were demonstrated when populations were disaggregated by country of origin, particularly for Mexican women (aOR 1.55, 95% CI 1.51–1.60; $p < 0.01$). Disparities worsened among both racial and country of origin subgroups in women with triple-negative disease.

Conclusion. Later breast cancer stage at presentation was observed among Hispanic populations when disaggregated by racial subgroup and country of origin. Socioeconomic disparities, as well as uncaptured disparities in access and/ or differential care, may drive these observed differences. Future studies with disaggregated data are needed to characterize outcomes in Hispanic communities and develop targeted interventions.

Breast cancer is the most commonly diagnosed cancer and the second leading cause of cancer mortality in women in the United States (US).[1] Prior research has examined the

factors influencing breast cancer presentation and outcome, including age, hormone receptor (HR) status, socioeconomic status, and race/ethnicity.[2–5] Among racial and ethnic subgroups, Hispanic and Black patients are disproportionately burdened with delayed diagnosis and poorer prognosis in breast cancer.[2]

Hispanic populations represent the second largest ethnic group in the US after non-Hispanic White individuals, and over 21% of Americans will identify as Hispanic by 2030.[6,7] Among Hispanic Americans, cancer is the leading cause of death, and breast cancer is the leading cause of cancer-specific death among Hispanic women.[8] Hispanic patients face significant barriers to healthcare access due to multiple social determinants of health, including structural discrimination, lower socioeconomic status, and lower health literacy.[9–11] In addition, Hispanic individuals have the highest uninsured rate in the US.[12] Among ethnically Hispanic patients, both Hispanic Black and Hispanic White patients are at higher risk of advanced-stage breast cancer when compared with non-Hispanic White patients.[13]

Many studies treat Hispanic patients as a monolithic group, and prior research aimed at stratifying in the context of breast cancer have mainly disaggregated by race.[4,14] While previous studies have suggested Black patients are at greater risk for triple-negative disease, there are limited studies examining this in Hispanic racial subgroups.[3,14,15] Furthermore, few studies have examined the influence of Hispanic country of origin in cancer outcomes.[4]

Given diverse cultural, socioeconomic, and immigration experiences of Hispanic subgroups, we hypothesize that there may be disparities in breast cancer stage at presentation on disaggregation. As breast cancer survival continues to improve due to advances in early detection and treatment, our study uses national data to evaluate stage at presentation with subgroup analysis by both receptor status and intersectional racial/ethnic subgroups, with further disaggregation by country of origin.[16] Our study encourages practicing oncologists to explore the diversity within Hispanic populations and analyze the social determinants of health unique to various subgroups. In analyzing variations in outcome, we hope to help physicians develop more targeted interventions to prevent delays in diagnosis for their Hispanic patients.

METHODS AND MATERIALS

Throughout this paper, we use the term 'Hispanic' in reference to individuals identified as belonging to the ethnic population with cultural lineage from Spain, Latin America (excluding Brazil), or the Caribbean; this definition is driven by how the National Cancer Database (NCDB) categorizes Hispanic populations; however, we

also reference literature on Latino populations given its relevance to this research. Hispanic populations are often mislabeled as a racial group rather than an ethnicity composed of various racial subgroups. References to race throughout this paper describe racial subgroups (non-Hispanic Black, Hispanic White, etc.), not the Hispanic population overall.

Data Source and Study Population

Data from 2004 to 2017 were extracted from the NCDB. The NCDB provides > 34 million records from hospital cancer registries, with data sourced from > 1500 cancer programs in both the continental US and Puerto Rico, accounting for approximately 70% of new cancer diagnoses.[17]

Women with breast cancer were included if information regarding their stage (0–IV) and Hispanic ethnicity were both available. They were stratified by race and categorized as non-Hispanic White, non-Hispanic Black, non-Hispanic other, Hispanic White, Hispanic Black, and Hispanic other. 'Other' categories included populations who identified as American Indian, Asian, Native Hawaiian, Pacific Islander, or other; individuals with unknown race were not included in this study. The Spanish/Hispanic origin variable was used to further disaggregate individuals identified as Hispanic into Mexican (including Chicano), Puerto Rican, Cuban, South or Central American (except Brazil), other specified Spanish/Hispanic Origin (includes Europe but excludes Dominican Republic), Spanish surname only, and Dominican (Republic). Patients missing data on stage, ethnicity, or other covariates included in our models were excluded.

Clinical and Sociodemographic Covariates

The dependent variable was breast cancer stage at presentation, and the primary independent variable was Hispanic ethnicity disaggregated by both race and country of origin. All models adjusted for additional independent variables, including age, facility type, year of diagnosis, and Charlson–Deyo comorbidity.

Given complexities in interpreting models that examine inequalities by race/ethnicity as well as socioeconomic status, epidemiological experts have encouraged a more thoughtful approach to covariate adjustment.[18,19] In this study, we use an interpretation of the effects of race/ethnicity proposed by Vanderweele and Robinson: "the health inequality that would remain for persons" after covariates had been set equal between populations.[20] As social determinants of health are likely mediators of interactions between dependent variables and stage at presentation, we also created models that further adjusted for three

variables: median household income in patient's area of residence (neighborhood income: < \$38,000, \$38,000–\$47,999, \$48,000–\$62,999, ≥ \$63,000), percentage of adults in the patient ZIP code who did not graduate from high school (neighborhood education: ≥ 21.0%, 13.0–20.9%, 7.0–12.9%, < 7.0%), and insurance status.

Statistical Analysis

Chi-square tests analyzed the distribution of categorical variables by racial subgroup. Adjusted odds ratios (aOR) with 95% confidence intervals (CIs) were created using ordinal logistic regression. Two models were run to assess the odds of presenting at a later stage, with one model disaggregated by race and the other disaggregated by country of origin. Each model adjusted for (1) age, facility type, year of diagnosis, and Charlson–Deyo comorbidity; and (2) previous covariates plus neighborhood income, neighborhood education, and insurance status. The population was then stratified by receptor status (HR+ with HER2−, any estrogen receptor/progesterone receptor status with HER2+, and triple negative), and the analysis was repeated. Non-Hispanic White women were used as the reference group in all models given that they were the largest population, consistent with prior Hispanic disaggregation studies.[4,13,21]

Analyses were performed using Stata/MP 17.0 (Stata-Corp LLC, College Station, TX, USA). This study was deemed exempt from the hospital Institutional Review Board given use of de-identified data.

RESULTS

Baseline Characteristics

Overall, 2,282,691 women were included, of whom 475,749 (20.84%) had stage 0 disease, 968,698 (42.44%) had stage 1 disease, 559,277 (24.50%) had stage 2 disease, 186,588 (8.17%) had stage 3 disease, and 92,379 (4.05%) had stage 4 disease. Patient characteristics are presented in Table 1.

Non-Hispanic Black and Hispanic Black patients had higher comorbidity rates, with 77.11% and 79.77% reporting a comorbidity score of 0, respectively (vs. 84.55% in non-Hispanic White patients). Non-Hispanic White patients were less likely to live in neighborhoods with low education: 13.09% non-Hispanic White compared with 57.32% Hispanic Black women. Non-Hispanic White and non-Hispanic Other populations were least likely (11.98% and 8.27%, respectively) and non-Hispanic Black women (41.55%) were most likely to live in the lowest-income neighborhoods. Hispanic White and Hispanic

Other populations were most likely to be uninsured (7.77% and 7.68%, respectively) compared with non-Hispanic White patients (1.18%). Hispanic patients were more likely to have Medicaid, with Hispanic Black populations reporting the highest proportion at 28.13%.

Breast Cancer Stage at Presentation

Compared with non-Hispanic White patients, Hispanic women overall had greater odds of presenting with later-stage breast cancer (aOR 1.19, 95% CI 1.18–1.21; $p < 0.01$) [electronic supplementary material (ESM) Table 1], with no significant differences noted after adjusting for socioeconomic status (aOR 1.00, 95% CI 0.99–1.02; $p = 0.39$). By racial subgroup stratification, all groups besides non-Hispanic Other had greater odds of presenting at a later stage than non-Hispanic White women (non-Hispanic Black: aOR 1.29, 95% CI 1.28–1.30, $p < 0.01$; Hispanic White: aOR 1.20, 95% CI 1.19–1.21, $p < 0.01$; Hispanic Black: aOR 1.11, 95% CI 1.03–1.19, $p \leq 0.01$; Hispanic Other: aOR 1.14, 95% CI 1.09–1.19, $p < 0.01$) (Fig. 1a; Table 2). Non-Hispanic Other women were less likely to present at a later stage (aOR 0.93, 95% CI 0.92–0.94; $p < 0.01$). After adjusting for socioeconomic status, odds to present at a later stage remained robust in non-Hispanic Black women (aOR 1.14, 95% CI 1.13–1.15; $p < 0.01$), while Hispanic Black women were found to have decreased odds of a later stage (aOR 0.88, 95% CI 0.82–0.95; $p < 0.01$).
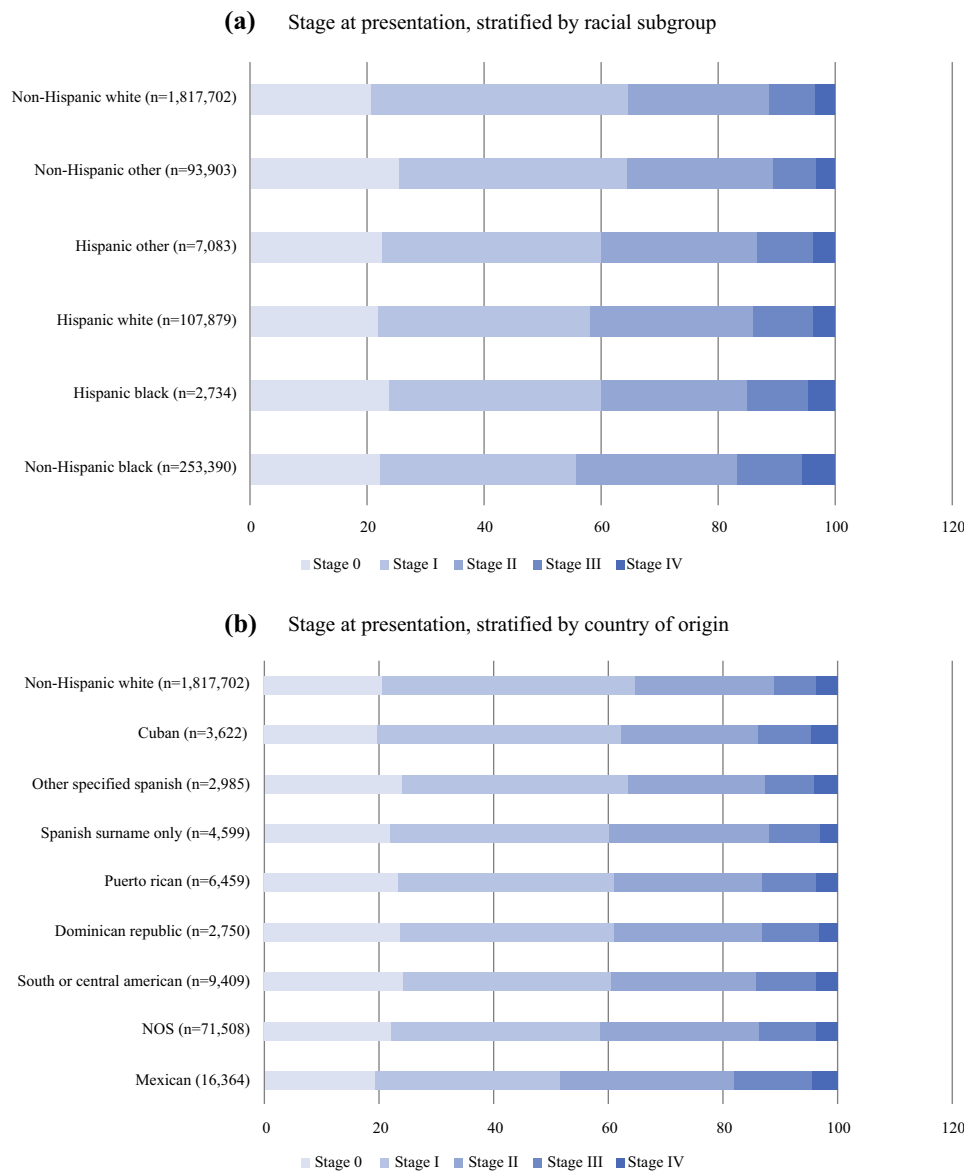
All subgroups by country of origin were found to have increased odds of presenting at a later stage compared with the reference non-Hispanic White group, except Puerto Rican (aOR 1.05, 95% CI 1.00–1.10; $p = 0.06$), Other Specified Spanish (aOR 0.98, 95% CI 0.92–1.05; $p = 0.65$), and Dominican (aOR 1.06, 95% CI 0.99–1.14; $p = 0.09$) women (Mexican: aOR 1.55, 95% CI 1.51–1.60, $p < 0.01$; Cuban: aOR 1.12, 95% CI 1.05–1.19, $p < 0.01$; South or Central American: aOR 1.09, 95% CI 1.05–1.13, $p < 0.01$; Not Otherwise Specified (NOS): aOR 1.18, 95% CI 1.16–1.19, $p < 0.01$; Spanish surname only: aOR 1.10, 95% CI 1.5–1.16, $p < 0.01$) (Fig. 1b; Table 2). When accounting for socioeconomic factors, odds of presenting at a later stage remained robust for Mexican women (aOR 1.22, 95% CI 1.18–1.25; $p < 0.01$) but decreased for other groups (Puerto Rican: aOR 0.89, 95% CI 0.85–0.93, $p < 0.01$; South or Central American: aOR 0.90, 95% CI 0.87–0.94, $p < 0.01$; Other Specified Spanish: aOR 0.85, 95% CI 0.79–0.90, $p < 0.01$; and Dominican populations: aOR 0.81, 95% CI 0.75–0.87; $p < 0.01$). Across all models, odds of later-stage disease appeared to decline over time.

Analyses were repeated in three subpopulations based on receptor status (Table 2; ESM Table 2). Disparities by race and country of origin appeared consistent or even

TABLE 1 Sample characteristics of the study population

Total sample: 2,282,691	Non-Hispanic			Hispanic		
	Non-Hispanic White (n = 1,817,702)	Non-Hispanic Black (n = 253,390)	Non-Hispanic other (n = 93,903)	Hispanic White (n = 107,879)	Hispanic Black (n = 2734)	Hispanic other (n = 7083)
Receptor status—N (%)	< **0.001**					
ER/PR +, HER2-	676,258 (37.20%)	79,372 (31.32%)	35,171 (37.45%)	38,949 (36.10%)	1015 (37.13%)	2958 (41.76%)
Any ER/PR status, HER2 +	125,929 (6.93%)	21,628 (8.54%)	9182 (9.78%)	10,103 (9.37%)	273 (9.99%)	756 (10.67%)
Triple negative	91,774 (5.05%)	28,878 (11.40%)	4896 (5.21%)	6806 (6.31%)	274 (10.02%)	515 (7.27%)
Unknown	923,741 (50.82%)	123, 512 (48.74%)	44,654 (47.55%)	52,021 (48.22%)	1172 (42.87%)	2854 (40.29%)
Age—median (25–75%)	63 (53-72)	60 (51-69)	57 (49-67)	57 (49-67)	59 (51-68)	57 (49-66)
Facility type—N (%)	< **0.001**					
Community cancer program	181,573 (9.99%)	18,696 (7.38%)	9928 (10.57%)	9504 (8.81%)	318 (11.63%)	648 (9.15%)
Comprehensive community cancer program	873,509 (48.06%)	93,916 (37.06%)	37,362 (39.79%)	43,775 (40.58%)	460 (16.83%)	2276 (32.13%)
Academic/research program	507,237 (27.91%)	104,503 (41.24%)	36,617 (38.99%)	40,043 (37.12%)	1274 (46.60%)	3428 (48.40%)
Integrated network cancer program	255,383 (14.05%)	36,275 (14.32%)	9996 (10.65%)	14,557 (13.49%)	682 (24.95%)	731 (10.32%)
Charlson–Deyo comorbidity coefficient—N (%)	< **0.001**					
Score of 0	1,536,857 (84.55%)	195,392 (77.11%)	80,952 (86.21%)	89,679 (83.13%)	2181 (79.77%)	5807 (81.99%)
Score of 1	219,709 (12.09%)	43,455 (17.15%)	10,706 (11.40%)	14,813 (13.73%)	435 (15.91%)	1013 (14.30%)
Score of 2	45,109 (2.48%)	9910 (3.91%)	1619 (1.72%)	2426 (2.25%)	87 (3.18%)	191 (2.70%)
Score of 3 or higher	16,027 (0.88%)	4633 (1.83%)	626 (0.67%)	961 (0.89%)	31 (1.13%)	72 (1.02%)
Percent no high school degree quartile—N (%)	< **0.001**					
> 17.6%	237,969 (13.09%)	94,912 (37.46%)	20,174 (21.48%)	53,776 (49.85%)	1567 (57.32%)	2828 (39.93%)
10.9–17.5%	421,301 (23.18%)	82,358 (32.50%)	18,602 (19.81%)	22,209 (20.59%)	571 (20.89%)	1735 (24.50%)
6.3–10.8%	562,747 (30.96%)	51,288 (20.24%)	25,707 (27.38%)	18,745 (17.38%)	372 (13.61%)	1489 (21.02%)
< 6.3%	595,685 (32.77%)	24,832 (9.80%)	29,420 (31.33%)	13,149 (12.19%)	224 (8.19%)	1031 (14.56%)
Median income quartile— N (%)	< **0.001**					
< $40,227	217,745 (11.98%)	105,273 (41.55%)	7767 (8.27%)	25,482 (23.62%)	986 (36.06%)	1659 (23.42%)
$40,227–$50,353	364,076 (20.03%)	52,983 (20.91%)	11,275 (12.01%)	23,327 (21.62%)	653 (23.88%)	1392 (19.65%)
$50,354–$63,332	435,926 (23.98%)	42,007 (16.58%)	18,334 (19.52%)	26,414 (24.48%)	487 (17.81%)	1595 (22.52%)
> $63,333	799,955 (44.01%)	53,127 (20.97%)	56,527 (60.20%)	32,656 (30.27%)	608 (22.24%)	2437 (34.41%)
Insurance status—N (%)	< **0.001**					
Uninsured	21,495 (1.18%)	8731 (3.45%)	2894 (3.08%)	8383 (7.77%)	101 (3.69%)	543 (7.68%)
Private insurance/managed care	956,003 (52.59%)	120,359 (47.50%)	56,330 (59.99%)	51,335 (47.59%)	1068 (39.06%)	3455 (48.78%)
Medicaid	67,219 (3.70%)	29,745 (11.74%)	10,421 (11.10%)	18,357 (17.02%)	769 (28.13%)	1250 (17.65%)
Medicare	756,814(41.64%)	91,317 (36.04%)	22,599 (24.07%)	28,916 (26.80%)	779 (28.49%)	1734 (24.48%)
Other government	16,171 (0.89%)	3238 (1.28%)	1659 (1.77%)	888 (0.82%)	17 (0.62%)	101 (1.43%)

ER estrogen receptor, *PR* progesterone receptor, *HER2* human epidermal growth factor receptor

(a)   Stage at presentation, stratified by racial subgroup

(b)   Stage at presentation, stratified by country of origin

FIG. 1 Proportion of individuals presenting with breast cancer stages 0–IV, stratified by (**a**) racial subgroup and (**b**) country of origin

larger in magnitude. Among patients with triple-negative disease, odds of delayed diagnosis worsened in all racial subgroups besides Hispanic Black (non-Hispanic Black: aOR 1.42, 95% CI 1.38–1.45, $p < 0.01$; non-Hispanic Other: aOR 1.09, 95% CI 1.03–1.15, $p < 0.01$; Hispanic White: aOR 1.32, 95% CI 1.26–1.38, $p < 0.01$; Hispanic Black: aOR 1.23, 95% CI 0.99–1.54, $p = 0.06$; Hispanic Other: aOR 1.35, 95% CI 1.16–1.59, $p < 0.01$) (Fig. 1a; Table 2). Odds similarly worsened for Mexican (aOR 1.59, 95% CI 1.43–1.76; $p < 0.01$), South or Central American (aOR 1.33, 95% CI 1.14–1.55; $p < 0.01$), and NOS (aOR 1.32, 95% CI 1.25–1.39; $p < 0.01$) women.

Subgroup analysis was conducted in Black women, with non-Hispanic Black used as the reference group (ESM

Tables 3 and 4). Hispanic Black women were less likely to present at later stage (aOR 0.89, 95% CI 0.83–0.95; $p < 0.01$), with the difference becoming stronger in the mediator analysis (aOR 0.81, 95% CI 0.76–0.87; $p < 0.01$).

DISCUSSION

This national study of over 2 million women with breast cancer found overall later stage at diagnosis in the aggregate Hispanic population, with further disparities in late presentation among specific Hispanic subgroups. Hispanic women (of Black or White race) were more likely to present at a later stage of cancer than non-Hispanic White

TABLE 2 Adjusted odds ratios with 95% confidence intervals comparing the odds of presenting at progressively later-stage disease for Hispanic women disaggregated by racial subgroup and Hispanic country of origin, with non-Hispanic White as the reference group

| | Total population (n = 2,282,691) | | | |
| | Adjusted | | Mediator analysis | |
	AOR (CI)	p value	AOR (CI)	p value
Race/ethnicity—N (%)				
Non-Hispanic White	1,817,702 (79.63%)	(reference)		
Non-Hispanic Black	253,390 (11.10%)	**1.29 (1.28–1.30)**	**< 0.01**	**1.14 (1.13–1.15)**	**< 0.01**
Non-Hispanic other	93,903 (4.11%)	**0.93 (0.92–0.94)**	**< 0.01**	**0.90 (0.89–0.91)**	**< 0.01**
Hispanic White	107,879 (4.73%)	**1.20 (1.19–1.21)**	**< 0.01**	1.01 (1.00–1.02)	0.07
Hispanic Black	2734 (0.12%)	**1.11 (1.03–1.19)**	**< 0.01**	**0.88 (0.82–0.95)**	**< 0.01**
Hispanic other	7083 (0.31%)	**1.14 (1.09–1.19)**	**< 0.01**	0.96 (0.92–1.01)	0.09
Hispanic country of origin—N (%)				
Mexican	16,364 (0.72%)	**1.55 (1.51–1.60)**	**< 0.01**	**1.22 (1.18–1.25)**	**< 0.01**
Puerto Rican	6459 (0.28%)	1.05 (1.00–1.10)	0.06	**0.89 (0.85–0.93)**	**< 0.01**
Cuban	3622 (0.16%)	**1.12 (1.05–1.19)**	**< 0.01**	1.01 (0.95–1.07)	0.79
South or Central American	9409 (0.41%)	**1.09 (1.05–1.13)**	**< 0.01**	**0.90 (0.87–0.94)**	**< 0.01**
Other Specified Spanish	2985 (0.13%)	0.98 (0.92–1.05)	0.65	**0.85 (0.79–0.90)**	**< 0.01**
Not Otherwise Specified	71,508 (3.13%)	**1.18 (1.16–1.19)**	**< 0.01**	1.00 (0.98–1.01)	0.68
Spanish Surname only	4599 (0.20%)	**1.10 (1.05–1.16)**	**< 0.01**	1.04 (0.99–1.10)	0.11
Dominican (Republic)	2750 (0.12%)	1.06 (0.99–1.14)	0.09	**0.81 (0.75–0.87)**	**< 0.01**

TABLE 2 continued

| | | Hormone receptor positive/HER2 negative (n = 833,723) | | | | Any hormone receptor status/HER2 positive (n = 167,881) | | | | Triple negative disease (n = 133,143) | | | |
| | | Adjusted | | Mediator analysis | | Adjusted | | Mediator analysis | | Adjusted | | Mediator analysis | |
		AOR (CI)	p value	AOR (CI)	p value	AOR (CI)	p value	AOR (CI)	p value	AOR (CI)	p value	AOR (CI)	p value
Race/ethnicity—N (%)													
Non-Hispanic White	1,817,702 (79.63%)	(Reference)				(Reference)				(Reference)			
Non-Hispanic Black	253,390 (11.10%)	**1.43 (1.41–1.45)**	**< 0.01**	**1.29 (1.27–1.31)**	**< 0.01**	**1.40 (1.36–1.43)**	**< 0.01**	**1.25 (1.22–1.29)**	**< 0.01**	**1.42 (1.38–1.45)**	**< 0.01**	**1.27 (1.23–1.30)**	**< 0.01**
Non-Hispanic other	93,903 (4.11%)	**1.09 (1.06–1.11)**	**< 0.01**	**1.04 (1.02–1.06)**	**< 0.01**	**1.07 (1.03–1.12)**	**< 0.01**	1.01 (0.97–1.05)	0.5	**1.09 (1.03–1.15)**	**< 0.01**	1.04 (0.99–1.10)	0.14
Hispanic White	107,879 (4.73%)	**1.29 (1.27–1.32)**	**< 0.01**	**1.10 (1.08–1.13)**	**< 0.01**	**1.27 (1.23–1.32)**	**< 0.01**	**1.06 (1.02–1.10)**	**< 0.01**	**1.32 (1.26–1.38)**	**< 0.01**	**1.09 (1.04–1.15)**	**< 0.01**
Hispanic Black	2734 (0.12%)	**1.27 (1.13–1.43)**	**< 0.01**	1.03 (0.92–1.16)	0.6	**1.31 (1.06–1.62)**	**0.016**	1.08 (0.87–1.34)	0.47	1.23 (0.99–1.54)	0.06	0.98 (0.79–1.23)	0.89
Hispanic other	7083 (0.31%)	**1.21 (1.13–1.30)**	**< 0.01**	1.04 (0.97–1.11)	0.28	**1.18 (1.04–1.35)**	**0.01**	0.99 (0.86–1.13)	0.84	**1.35 (1.16–1.59)**	**< 0.01**	1.15 (0.98–1.34)	0.09
Hispanic country of origin—N (%)													
Mexican	16,364 (0.72%)	**1.64 (1.56–1.72)**	**< 0.01**	**1.32 (1.26–1.38)**	**< 0.01**	**1.44 (1.32–1.56)**	**< 0.01**	**1.11 (1.02–1.21)**	**0.02**	**1.59 (1.43–1.76)**	**< 0.01**	**1.23 (1.11–1.37)**	**< 0.01**
Puerto Rican	6459 (0.28%)	**1.16 (1.07–1.25)**	**< 0.01**	1.01 (0.93–1.09)	0.84	**1.20 (1.03–1.39)**	**0.02**	1.00 (0.86–1.16)	0.99	1.13 (0.95–1.34)	0.16	1.00 (0.84–1.18)	0.97
Cuban	3622 (0.16%)	**1.20 (1.07–1.34)**	**< 0.01**	1.08 (0.96–1.21)	0.18	**1.27 (1.02–1.57)**	**0.03**	1.11 (0.89–1.38)	0.35	1.31 (0.98–1.73)	0.07	1.19 (0.90–1.58)	0.23
South or Central American	9409 (0.41%)	**1.25 (1.17–1.33)**	**< 0.01**	1.05 (0.98–1.12)	0.19	**1.19 (1.06–1.34)**	**< 0.01**	0.96 (0.86–1.09)	0.55	**1.33 (1.14–1.55)**	**< 0.01**	1.05 (0.89–1.22)	0.58
Other Specified Spanish	2985 (0.13%)	0.98 (0.87–1.12)	0.80	**0.88 (0.77–1.00)**	**< 0.05**	1.02 (0.80–1.29)	0.88	0.90 (0.71–1.14)	0.38	0.95 (0.70–1.28)	0.72	0.84 (0.62–1.14)	0.27
Not Otherwise Specified	71,508 (3.13%)	**1.26 (1.23–1.29)**	**< 0.01**	**1.08 (1.05–1.11)**	**< 0.01**	**1.27 (1.22–1.33)**	**< 0.01**	**1.07 (1.02–1.12)**	**< 0.01**	**1.32 (1.25–1.39)**	**< 0.01**	**1.10 (1.04–1.16)**	**< 0.01**
Spanish Surname only	4599 (0.20%)	**1.17 (1.06–1.29)**	**< 0.01**	**1.10 (1.00–1.22)**	**< 0.05**	1.02 (0.83–1.24)	0.86	0.96 (0.78–1.17)	0.67	**1.10 (0.87–1.38)**	**< 0.01**	1.02 (0.81–1.28)	0.87
Dominican (Republic)	2750 (0.12%)	**1.16 (1.03–1.30)**	**0.01**	0.90 (0.80–1.02)	0.09	1.17 (0.95–1.44)	0.13	0.89 (0.72–1.10)	0.27	1.10 (0.87–1.38)	0.72	0.72	0.72

Analysis is stratified by hormone and HER2 receptor status: (a) hormone receptor-positive, HER2-negative; (b) any hormone receptor status, HER2-positive; and (c) triple negative. Initial models adjusted for age, year of diagnosis, facility type, and Charlson–Deyo comorbidity coefficient. Mediator analysis models further adjusted for median household income for each patient's area of residence, percentage of adults in the patient's ZIP code who did not graduate from high school, and insurance status. Values that are statistically significant with a prespecified alpha of 0.05 are shown in bold

aOR adjusted odds ratio, *CI* confidence interval

women. Additionally, Mexican, Cuban, and South or Central American women were also more likely to present with later-stage disease. Disparities in stage remained for Mexican women after socioeconomic adjustment. Our findings suggest that social determinants of health contribute to why certain Hispanic groups present at a later stage, but cannot fully explain late-stage presentation seen in Mexican women. As Hispanic women were more likely to live in neighborhoods with low income and education, our findings support efforts to target social determinants of health as a means to mitigate breast cancer disparities.

Healthcare disparities among Hispanic patients have been driven by a number of social determinants of health, including structural racism, lower socioeconomic status, language barriers, and lack of insurance.[9,12,22] Disparities persist even between Hispanic subgroups, with noted differences in insurance status and healthcare utilization.[23] These factors appear to be particularly meaningful for Cuban and Central or South American populations in explaining their increased risk of presenting with later-stage breast cancer. For patients of Cuban descent in our study, there were no significant differences when compared with non-Hispanic White populations after we adjusted for socioeconomic social determinants of health. For patients of Mexican descent, lower rates of breast cancer screening could play a significant role. Ramirez et al. found that Mexican women along the Texas-Mexican border had the lower rates of mammograms/clinical breast examinations when compared with Central American women in San Francisco, Cuban women in Miami, and Puerto Rican women in New York.[24]

While 'Central or South American' remains a substantially heterogenous population, disruptions in medical care due to recent migrations may mediate the increased risk. According to the United Nations Population Division, nearly 17.6 million South Americans lived outside the country of their birth. While South American immigrants tend to be slightly more educated and more likely to participate in the labor force, political and humanitarian crises may have led to higher rates of mobility and lower rates of preventative screening prior to arrival in the US.[25,26] Given more recent immigrations compared with other Hispanic populations, language barriers may be more burdensome in this population. It is important to emphasize that the Central or South American population merit further disaggregation to provide meaningful explanations of observed differences given the inherent diversity in this large group.

Aside from experiences prior to entering the US, diverse immigration patterns have resulted in variations in geographic settlement among Hispanic subgroups, further contributing to differences in health status.[27] Citizenship may further play a role, particularly with regard to healthcare access. In our study, Puerto Rican women did not have delayed diagnosis when compared with non-Hispanic White populations. As US citizens, Puerto Rican populations may have better access to healthcare resources and Medicaid than other Hispanic populations.

Our study adds to the growing body of literature supporting a more granular approach in identifying groups for targeted interventions. Research efforts should be directed towards (1) greater understanding of the drivers of disparities through quantitative and qualitative studies; and (2) piloting targeted interventions to benefit the most vulnerable. Additionally, understanding how social factors can mediate disparities in stage at presentation for disaggregated Hispanic patients may inform our understanding of other groups that benefit from disaggregation, such as the Asian American, Native Hawaiian, and Pacific Islander populations.[28,29]

A challenge in disaggregating Latino data comes from the complex cultural heritages resulting from colonization in both North and South America. Many Latinos today carry a combination of European, Native American, and/or African ancestry based on Latin American country of origin.[30,31] Furthermore, differential assimilation to American language and culture by Hispanic populations may also drive variability in access and acceptance to cancer care, including differences in language barriers, socioeconomic status, nativity status, and healthcare literacy, as well as barriers that result from the deleterious effects of structural racism.[31–34]

At the intersectionality of race and ethnicity, our study shows that both Hispanic Black and non-Hispanic Black women are more likely to present with later-stage disease. Although adjusting for income/education ameliorated in Hispanic Black populations, non-Hispanic Black women remained at increased odds. When further analyzing by receptor status, Hispanic Black women were less likely to have triple-negative disease than non-Hispanic Black women. This is consistent with previous studies showing higher rates of triple negative disease in non-Hispanic Black women when compared with Hispanic Black women.[3] Few studies have explored the intersecting roles that race and ethnicity play in influencing health behaviors and outcome. LaVeist-Ramos et al. observed that ethnicity plays a dominant role in Hispanic Black communities for health behaviors (diet, alcohol use, smoking status), while race plays a more prominent role in health services (insurance coverage and last visit to the doctor).[35] Our findings reinforce the intersectionality inherent in identity, which when analyzed in the context of the social determinants of health, collectively contribute to outcome disparities.

Limitations

Limitations include those related to the primary dataset. Although the NCDB contains the majority of cancer cases in the US, it may not be representative of cancer in the population as a whole. Additionally, Hispanic subgroups defined by NCDB are imperfect, with categories such as Central and South American grouping a number of communities together; certain non-descript groupings such as Other Spanish or NOS are also difficult to interpret.[36] Grouping of Hispanic populations by race and country of origin may miss other important cultural factors that may play a role in health disparities, such as language or cross-cultural identity. We also acknowledge that large population studies can miss the vast heterogeneity within cultural identity groups and recognize that an individual patient's identity goes beyond their cultural and/or ethnic groups. Certain groupings created due to small sample size, such as 'Hispanic Other', may obscure potential disparities within traditionally underrepresented groups, such as Indigenous populations. Missingness in the data should also be acknowledged, particularly with regard to patient receptor status; nearly 40–50% of patients were missing information on the receptor status in each of the subgroups. Lastly, we recognize the imperfect nature of socioeconomic corrections given neighborhood-level data on income and education available in NCDB, as uncaptured disparities due to structural racism may not be adjusted for accurately.

Future Study

Our findings support several areas of future exploration. While this work suggests that socioeconomic status may mediate some of the observed disparities among Hispanic populations, further studies with more granular patient-level data are needed to validate this observation. Additionally, some groups, such as Mexican populations, continued to have higher odds of presenting at later-stage cancer in spite of adjustment for socioeconomic status, suggesting other mechanisms may be at play. Future studies should also explore language, social ties, cultural factors, and differential participation in cancer screening that may account for some of the observed disparities. Furthermore, studies should explore the role of acculturation and immigration in healthcare access and stage at presentation; differences may be observed between newly arrived immigrants and other immigrant populations who spent prolonged periods of time in the US. Disparities may further exist due to differences in geographic settlement of Hispanic subpopulations and access to targeted resources, which merit further exploration. Additionally, differences between Hispanic Black and non-Hispanic Black populations remain relatively understudied. Despite significant genetic admixture among Hispanic populations, future research could also consider exploring any potential differences in prevalence of known genetic mutations between Hispanic subgroups. Given their aggregation in this study, populations such as Central and South Americans should be better examined in future work. Importantly, qualitative studies are necessary to shed further light on the reasons underlying the observed disparities, and may help to clarify modifiable factors that could be leveraged to promote equity. From a treatment perspective, future studies should also consider access to treatment in these populations, with particular focus on any differential access to breast conservation versus mastectomy with or without reconstruction. Further work is needed to identify the causes of these disparities and may take the form of qualitative research.

CONCLUSION

There are significant disparities in breast cancer stage at presentation across Hispanic populations stratified by race and country of origin, with socioeconomic status potentially playing a large role in the observed disparities. Further research on potential mechanisms, including language, environmental, and cultural factors, should be explored to guide development of targeted interventions to improve both economic and access inequities, which could lead to improved outcomes in vulnerable Hispanic subpopulations.

REFERENCES

1. Sung H, Ferlay J, Siegel RL, et al. Global Cancer Statistics 2020: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries. *CA Cancer J Clin*. 2021;71(3):209–49. https://doi.org/10.3322/caac.21660.

2. Lantz PM, Mujahid M, Schwartz K, et al. The influence of race, ethnicity, and individual socioeconomic factors on breast cancer stage at diagnosis. *Am J Public Health*. 2006;96(12):2173–8. https://doi.org/10.2105/AJPH.2005.072132.

3. Champion CD, Thomas SM, Plichta JK, et al. Disparities at the intersection of race and ethnicity: examining trends and outcomes in hispanic women with breast cancer. *JCO Oncol Pract*. 2022;18(5):e827–38. https://doi.org/10.1200/OP.20.00381.

4. Ooi SL, Martinez ME, Li CI. Disparities in breast cancer characteristics and outcomes by race/ethnicity. *Breast Cancer Res Treat*. 2011;127(3):729–38. https://doi.org/10.1007/s10549-010-1191-6.

5. El Saghir NS, Seoud M, Khalil MK, et al. Effects of young age at presentation on survival in breast cancer. *BMC Cancer*. 2006;6(1):194. https://doi.org/10.1186/1471-2407-6-194.

6. 3 ways that the U.S. population will change over the next decade. PBS NewsHour. Published 2 January 2020. https://www.pbs.org/newshour/nation/3-ways-that-the-u-s-population-will-change-over-the-next-decade. Accessed 14 Jan 2022.

7. US Census Bureau. 2019 Population Estimates by Age, Sex, Race and Hispanic Origin. Census.gov. https://www.census.gov/newsroom/press-kits/2020/population-estimates-detailed.html. Accessed 29 Oct 2021.

8. American Cancer Society. Cancer Facts & Figures for Hispanic and Latino People. American Cancer Society. https://www.cancer.org/research/cancer-facts-statistics/hispanics-latinos-facts-figures.html. Accessed 15 Nov 2021.

9. Escarce JJ, Kapur K. Access to and Quality of Health Care. In: Tienda M, Mitchell F (eds). Hispanics and the future of America. National Academies Press (US); 2006. https://www.ncbi.nlm.nih.gov/books/NBK19910/. Accessed 29 Oct 2021.

10. Becerra BJ, Arias D, Becerra MB. Low health literacy among immigrant hispanics. *J Racial Ethn Health Disparities*. 2017;4(3):480–3. https://doi.org/10.1007/s40615-016-0249-5.

11. Sheppard VB, Williams KP, Wang J, Shavers V, Mandelblatt JS. An examination of factors associated with healthcare discrimination in latina immigrants: the role of healthcare relationships and language. *J Natl Med Assoc*. 2014;106(1):15–22. https://doi.org/10.1016/S0027-9684(15)30066-3.

12. Keisler-Starkey K, Bunch LN. Health Insurance Coverage in the United States: 2019. US Census Bureau; 2020. pp. 60-271.

13. Banegas MP, Li CI. Breast cancer characteristics and outcomes among Hispanic Black and Hispanic White women. *Breast Cancer Res Treat*. 2012;134(3):1297–304. https://doi.org/10.1007/s10549-012-2142-1.

14. Li CI, Malone KE, Daling JR. Differences in breast cancer hormone receptor status and histology by race and ethnicity among women 50 years of age and older. *Cancer Epidemiol Biomark*. 2002;11(7):601–7.

15. Elledge RM, Clark GM, Chamness GC, Osborne CK. Tumor biologic factors and breast cancer prognosis among white, Hispanic, and black women in the United States. *J Natl Cancer Inst*. 1994;86(9):705–12. https://doi.org/10.1093/jnci/86.9.705.

16. DeSantis CE, Ma J, Gaudet MM, et al. Breast cancer statistics, 2019. *CA Cancer J Clin*. 2019;69(6):438–51. https://doi.org/10.3322/caac.21583.

17. About the National Cancer Database. American College of Surgeons. https://www.facs.org/quality-programs/cancer/ncdb/about. Accessed 29 Oct 2021.

18. Meghani SH, Chittams J. Controlling for socioeconomic status in pain disparities research: all-else-equal analysis when "all else" is not equal. *Pain Med Malden Mass*. 2015;16(12):2222–5. https://doi.org/10.1111/pme.12829.

19. Nuru-Jeter AM, Michaels EK, Thomas MD, Reeves AN, Thorpe RJ, LaVeist TA. Relative roles of race versus socioeconomic position in studies of health inequalities: a matter of interpretation. *Annu Rev Public Health*. 2018;39:169–88. https://doi.org/10.1146/annurev-publhealth-040617-014230.

20. VanderWeele TJ, Robinson WR. On the causal interpretation of race in regressions adjusting for confounding and mediating variables. *Epidemiol Camb Mass*. 2014;25(4):473–84. https://doi.org/10.1097/EDE.0000000000000105.

21. Siegel RL, Fedewa SA, Miller KD, et al. Cancer statistics for Hispanics/Latinos, 2015. *CA Cancer J Clin*. 2015;65(6):457–80. https://doi.org/10.3322/caac.21314.

22. Timmins CL. The impact of language barriers on the health care of Latinos in the United States: a review of the literature and guidelines for practice. *J Midwif Womens Health*. 2002;47(2):80–96. https://doi.org/10.1016/s1526-9523(02)00218-0.

23. Weinick RM, Jacobs EA, Stone LC, Ortega AN, Burstin H. Hispanic healthcare disparities: challenging the myth of a monolithic Hispanic population. *Med Care*. 2004;42(4):313–20. https://doi.org/10.1097/01.mlr.0000118705.27241.7c.

24. Ramirez AG, Talavera GA, Villarreal R, et al. Breast cancer screening in regional Hispanic populations. *Health Educ Res*. 2000;15(5):559–68. https://doi.org/10.1093/her/15.5.559.

25. Batalova JBJL and J. South American Immigrants in the United States. migrationpolicy.org. Published 15 February 2022. https://www.migrationpolicy.org/article/south-american-immigrants-united-states. Accessed 29 June 2022.

26. Batalova J, Babich E. Central American Immigrants in the United States. migrationpolicy.org. Published 6 August 2021. https://www.migrationpolicy.org/article/central-american-immigrants-united-states. Accessed 29 June 2022.

27. James G. Latinx Great Migrations—Migration History 1850-2017. University of Washington—The Great Depression in Washington State Project. https://depts.washington.edu/moving1/latinx_migration.shtml. Accessed 14 Jan 2022.

28. Dee EC, Chen S, Santos PMG, Wu SZ, Cheng I, Gomez SL. Anti-Asian American Racism: a wake-up call for population-based cancer research. *Cancer Epidemiol Biomark Prev*. 2021;30(8):1455–8. https://doi.org/10.1158/1055-9965.EPI-21-0445.

29. Jain B, Ng K, Santos PMG, et al. Prostate cancer disparities in risk group at presentation and access to treatment for Asian Americans, Native Hawaiians, and Pacific Islanders: a study with disaggregated ethnic groups. *JCO Oncol Pract*. 2022;18(1):e204–18. https://doi.org/10.1200/OP.21.00412.

30. Stern MC. Prostate cancer in US Latinos: what have we learned and where should we focus our attention. In: Ramirez AG, Trapido EJ (eds). *Advancing the science of cancer in Latinos*. Springer; 2020. pp. 57–67. https://doi.org/10.1007/978-3-030-29286-7_5

31. Stern MC, Fejerman L, Das R, et al. Variability in cancer risk and outcomes within US Latinos by national origin and genetic ancestry. *Curr Epidemiol Rep*. 2016;3:181–90. https://doi.org/10.1007/s40471-016-0083-7.

32. Andrulis DP, Brach C. Integrating literacy, culture, and language to improve health care quality for diverse populations. *Am J*

Health Behav. 2007;31(Suppl 1):S122–33. https://doi.org/10.5555/ajhb.2007.31.supp.S122.

33. Dominguez K, Penman-Aguilar A, Chang MH, et al. Vital signs: leading causes of death, prevalence of diseases and risk factors, and use of health services among Hispanics in the United States—2009–2013. *Morb Mortal Wkly Rep*. 2015;64(17):469–78.

34. Lavizzo-Mourey RJ, Besser RE, Williams DR. Understanding and mitigating health inequities—past, current, and future directions. NEJM. https://doi.org/10.1056/NEJMp2008628. Accessed 5 Jan 2022.

35. LaVeist-Ramos TA, Galarraga J, Thorpe RJ, Bell CN, Austin CJ. Are black Hispanics black or Hispanic? Exploring disparities at the intersection of race and ethnicity. *J Epidemiol Community Health*. 2012;66(7):e21. https://doi.org/10.1136/jech.2009.103879.

36. Racism is a Public Health Crisis. https://www.apha.org/topics-and-issues/health-equity/racism-and-health/racism-declarations. Accessed 3 Nov 2021.