Research Square

Preprints are preliminary reports that have not undergone peer review.
They should not be considered conclusive, used to inform clinical practice,
or referenced by the media as validated information.

Associations Between the 21-Gene Oncotype DX Recurrence Score, Ki67, and Race in Early Breast Cancer: An Analysis of the National Cancer Database

NITHYA Krishnamurthy
Icahn School of Medicine at Mount Sinai
Deukwoo Kwon
Icahn School of Medicine at Mount Sinai
Grace Van Hyfte
Icahn School of Medicine at Mount Sinai
Joseph Sparano
Tisch Cancer Institute
Amy Tiersten
Tisch Cancer Institute
Rima Patel
rima.patel2@mssm.edu

Tisch Cancer Institute

Research Article

Keywords: Oncotype, Race, Recurrence Score, Healthcare Disparities

Posted Date: February 27th, 2026

DOI: https://doi.org/10.21203/rs.3.rs-8397787/v1

License: This work is licensed under a Creative Commons Attribution 4.0 International License.
Read Full License

Additional Declarations: No competing interests reported.

Abstract

Purpose:
The study examines the association between the Oncotype DX recurrence score (RS) and Ki67 expression across racial groups using the National Cancer Database (NCDB) to understand racial disparities in hormone receptor (HR)-positive, HER2-negative early breast cancer (EBC).

Methods:
Women with HR-positive, HER2-negative EBC and 0–3 positive lymph nodes diagnosed in 2018–2019 in the NCDB were included. RS was categorized as low (0–10), intermediate (11–25), and high (26–100); Ki67 as low (≤5%), intermediate (6–29%), and high (≥30%). Wilcoxon and chi-square tests assessed differences; Fleiss' Kappa measured RS–Ki67 concordance by race.

Results:
Among 43,898 patients, 17% were node positive. The cohort was 78% Non-Hispanic White (NHW), 8% Non-Hispanic Black (NHB), 6% Hispanic, and 4% Asian American/Pacific Islander (AAPI). RS and Ki67 distributions differed significantly by race, with NHB patients showing the highest proportions of high RS and Ki67 ($p<0.0001$). Overall agreement was slight (Kappa=0.19, $p<0.0001$), with fair agreement for high RS and high Ki67 (Kappa=0.35, $p<0.0001$). Stratified by race, agreement remained slight for NHW, Hispanic, and AAPI patients, but was fair for NHB patients (Kappa=0.24, $p=0.002$). The strongest concordance between high RS and high Ki67 was seen in NHB patients (Kappa=0.39, $p<0.0001$).

Conclusions:
In this NCDB cohort, Ki67 and RS were only slightly concordant overall, with fair agreement observed among patients with high Ki67 and RS. The strongest agreement between Ki67 and RS was noted in the Black subgroup compared to other races, likely due to the higher proportion of patients with high Ki67 and RS in this subgroup.

Introduction

The 21-gene Oncotype DX recurrence score (RS) is a clinically validated genomic assay that offers both prognostic and predictive insights in patients with hormone receptor (HR)-positive, human epidermal growth factor receptor 2 (HER2)-negative early-stage breast cancer (EBC) [1–3]. Specifically, the RS provides critical information regarding the risk of distant recurrence and the likely benefit derived from adjuvant chemotherapy, thereby playing a pivotal role in guiding treatment decisions for this breast cancer subtype [4–5]. The landmark TAILORx and RxPONDER trials demonstrated that postmenopausal patients with HR-positive, HER2-negative EBC, 0–3 involved axillary lymph nodes and RS < 26 do not derive significant benefit from chemotherapy, whereas premenopausal patients with node-negative disease and RS between 16–25 or positive lymph nodes may experience benefit [6–7].

Alongside genomic assays, immunohistochemical markers such as Ki67, a nuclear protein associated with cellular proliferation, have been widely studied for their prognostic utility in HR-positive EBC [8–9].

Ki67 expression has been correlated with tumor aggressiveness and recurrence risk; however, its clinical implementation remains variable due to issues with standardization and interobserver variability [10–11]. The International Ki67 Working Group (IKWG) found that there was high concordance among pathologists for specimens with Ki67 ≤ 5% or ≥ 30% [12]. Based on this, the IKWG recommended classification of Ki67 by ≤ 5% (low), 6%–29% (intermediate), and ≥ 30% (high) and that the clinical utility of Ki67 may be limited to low and high groups and in patients with HR-positive, HER2-negative, breast tumors up to 5 centimeters and with up to three lymph nodes involved [10, 12].

The relationship between Ki67 and the RS remains uncertain. In a retrospective study of 525 patients with HR-positive, HER2-negative breast cancer, only slight overall agreement was observed between Ki67 and RS, with fair concordance limited to cases with high Ki67 (≥ 30%) and high RS (≥ 26), underscoring their limited interchangeability [13]. In a separate cohort study of HR-positive, HER2-negative breast cancer, a moderate correlation was observed between Ki67 and RS; notably, among patients with low genomic risk breast cancers, high Ki67 expression was associated with an increased risk of secondary endocrine resistance, suggesting that Ki67 may provide complementary prognostic information in select subgroups [14].

Racial disparities in breast cancer outcomes have been well-documented, particularly among patients with HR-positive disease [15–18]. Black patients with HR-positive EBC consistently experience poorer survival outcomes compared to White patients, influenced by both biological factors and structural inequities [17–23]. Black patients with ER-positive breast cancer had nearly double the recurrence risk of non-Hispanic White patients, and outcome disparity persisted even after adjusting for obesity, stage at diagnosis, and treatment adherence[24–25]. Black women have poorer overall survival even after adjusting for recurrence score and clinical factors, and this pattern appears in both node negative and node positive disease. Differences in standard pathological features, including lympho-vascular invasion and recurrence score distribution, do not explain the gap, suggesting contributions from biological factors not reflected in routine measures [19]. A recent pooled analysis of eight National Surgical Adjuvant Breast and Bowel Project (NSABP) trials involving over 9,700 patients with localized breast cancer found that Black race was independently associated with worse distant recurrence-free survival (DRFS) in HR-positive disease, even after adjusting for clinicopathologic variables such as age, tumor size, nodal status, body mass index, and treatment type [26]. These findings underscore the persistent and biologically nuanced impact of race on breast cancer outcomes, particularly in HR-positive subtypes.

Although the RS is widely used in clinical practice, a large SEER-Oncotype DX cohort study found that Black patients were more likely to have high-risk scores and experienced higher breast cancer-specific mortality compared to White patients with similar scores, indicating lower prognostic accuracy of the RS in Black race [27]. Nonetheless, the Oncotype DX RS remains predictive of chemotherapy benefit across racial/ethnic groups [28]. Other studies have reported racial differences in Ki67 but not in RS [29].

Despite extensive research on RS, Ki67, and race as individual prognostic factors in HR-positive, HER2-negative EBC, few studies have investigated how these variables interact. In particular, the effect of race

on the relationship between Ki67 and RS is not well understood. To address this gap, the present study utilized the National Cancer Database (NCDB) to evaluate the association between Ki67 expression and RS across racial groups in patients with HR-positive, HER2-negative EBC [30].

Methods

Patient population and data source

This study included patients aged 18 years and older diagnosed between 2018 and 2019 with HR-positive, HER2-negative EBC, characterized by tumors up to 5 cm in size (T1–T2) and up to three involved axillary lymph nodes (N0–N1). Eligible patients had documented information on the 21-gene Oncotype DX RS, IHC-measured Ki67 expression, self-reported race/ethnicity, and survival. Patients with metastatic disease were excluded. Data were obtained from the NCDB, a nationwide registry jointly maintained by the American College of Surgeons' Commission on Cancer and the American Cancer Society. The NCDB collects data from over 1,500 accredited cancer programs across the United States, representing more than 70% of newly diagnosed cancer cases. Certified tumor registrar's abstract data using standardized national protocols, capturing demographic, clinical, and treatment information. To ensure accuracy, at least 10% of medical records and registry abstracts are externally reviewed every three years at Commission-accredited institutions [31]. The 2019 NCDB dataset was used to identify the study cohort. This study was approved by the Mount Sinai Health System Institutional Review Board and conducted in accordance with the ethical principles outlined in the Declaration of Helsinki.

Study Design

Patients diagnosed between 2018 and 2019 with HR-positive, HER2-negative EBC and 0–3 involved lymph nodes were identified within the NCDB (Fig. 1). Although the initial NCDB extract spanned 2004–2019, exclusions for missing RS or Ki-67 left only cases from 2018–2019 in the final analytic cohort. RS was categorized into three risk groups: low (0–10), intermediate (11–25), and high (26–100) based on the TAILORx trial. Ki-67 values were grouped as low (≤ 5%), intermediate (6%–29%), and high (≥ 30%), based on the prognostic classification recommended by the International Ki67 Working Group [32].

Demographic data, including patient age, race/ethnicity, income, education, insurance status, care facility type, and geographic location, were collected. Neighborhood-level data on education and income were derived by matching patient ZIP codes to 2016 American Community Survey data (spanning 2012–2016), and categorized into quartiles. Education level reflected the proportion of adults ≥ 25 years without a high school diploma, and income was reported as median household income, adjusted for 2016 inflation.

Clinicopathologic features such as tumor grade, size, nodal involvement, and treatment characteristics (e.g., chemotherapy, endocrine therapy) were also extracted. Patients were stratified by race/ethnicity

into non-Hispanic White (NHW), non-Hispanic Black (NHB), Hispanic, Asian-American or Pacific Islander (AAPI), and other racial/ethnic groups.

Statistical Analyses

Patient demographics, tumor characteristics, and treatment variables were summarized by race/ethnicity. Categorical variables were compared using chi-square tests, and continuous variables were compared using the Wilcoxon rank-sum test. Concordance between Ki67 classification and RS risk categories based on racial subgroup was evaluated using the Fleiss' Kappa statistic, with corresponding p-values reported. Statistical significance was defined as $p < .05$. All statistical tests were two-sided, and analyses were performed using SAS version 9.4 and R, version 4.4.0.

Results

Patient Characteristics

Baseline demographic characteristics of the patient population, categorized and compared by race and ethnicity, are shown in Table 1. There were 43,898 eligible patients who were included in the final analysis (Fig. 1). Overall, 78.1% (N = 34275) of patients were NHW, 8.0% (N = 3502) were NHB, 6.4% (N = 2819) were Hispanic, 4.3% (N = 1870) were AAPI, and the remainder were American Indian and Alaska Native (AIAN), other or unknown (3.3%; N = 1432).

Table 1
Patient Baseline Demographic Characteristics

Patient Characteristics % (N)	**Overall 100% (43898)**	**Non-Hispanic White 78.08% (34275)**	**Non-Hispanic Black 7.98% (3502)**	**Hispanic 6.42% (2819)**	**AAPI 4.26% (1870)**	**AIAN/ Other/ Unknown 3.26% (1432)**	**P value**
Age							< 0.0001
Age < 50	18.78% (8244)	16.90% (5793)	20.82% (729)	28.59% (806)	33.48% (626)	20.25% (290)	
Age 50–59	25.50% (11195)	24.91% (8539)	26.87% (941)	29.19% (823)	27.22% (509)	26.75% (383)	
Age 60–69	34.59% (15185)	35.51% (12171)	34.75% (1217)	29.41% (829)	26.63% (498)	32.82% (470)	
Age ≥ 70	21.13% (9274)	22.68% (7772)	17.56% (615)	12.81% (361)	12.67% (237)	20.18% (289)	
Insurance							< 0.0001
Private	54.18% (23786)	54.19% (18574)	49.11% (1720)	52.75% (1487)	66.10% (1236)	53.70% (769)	
Medicaid	5.62% (2465)	4.11% (1410)	11.22% (393)	14.08% (397)	9.14% (171)	6.56% (94)	
Medicare	36.79% (16150)	38.98% (13361)	35.07% (1228)	24.65% (695)	20.75% (388)	33.38% (478)	
Not insured	1.25% (549)	0.70% (241)	1.74% (61)	6.81% (192)	1.55% (29)	1.82% (26)	
Other/Unknown	2.16% (948)	2.01% (689)	2.86% (100)	1.70% (48)	2.46% (46)	4.54% (65)	
Income							< 0.0001
<$38,000	10.49% (4604)	8.74% (2997)	25.96% (909)	16.46% (464)	3.74% (70)	11.45% (164)	
$38k-47,999	16.13% (7079)	16.16% (5540)	18.16% (636)	16.42% (463)	9.63% (180)	18.16% (260)	
$48k-62,999	22.84% (10027)	23.25% (7969)	19.45% (681)	24.41% (688)	20.16% (377)	21.79% (312)	
≥$63,000	36.99% (16236)	38.21% (13096)	21.22% (743)	29.51% (832)	55.72% (1042)	36.52% (523)	

Patient Characteristics % (N)	Overall 100% (43898)	Non-Hispanic White 78.08% (34275)	Non-Hispanic Black 7.98% (3502)	Hispanic 6.42% (2819)	AAPI 4.26% (1870)	AIAN/ Other/ Unknown 3.26% (1432)	P value
Unknown	13.56% (5952)	13.63% (4673)	15.22% (533)	13.20% (372)	10.75% (201)	12.08% (173)	
Education (% without High School degree)							< 0.0001
≥ 21%	11.42% (5011)	8.06% (2762)	23.16% (811)	35.86% (1011)	13.53% (253)	12.15% (174)	
13-20.9%	18.48% (8111)	17.63% (6044)	27.93% (978)	19.33% (545)	16.42% (307)	16.55% (237)	
7-12.9%	28.56% (12537)	30.20% (10350)	21.82% (764)	17.17% (484)	26.63% (498)	30.80% (441)	
<7%	28.01% (12297)	30.50% (10455)	11.91% (417)	14.44% (407)	32.67% (611)	28.42% (407)	
Unknown	13.54% (5942)	13.61% (4664)	15.19% (532)	13.20% (372)	10.75% (201)	12.08% (173)	
Region							< 0.0001
Urban/Metro	86.04% (37769)	84.20% (28859)	92.43% (3237)	94.79% (2672)	96.84% (1811)	83.10% (1190)	
Rural	1.23% (538)	1.43% (491)	0.74% (26)	0.18% (5)	0.11% (2)	0.98% (14)	
Unknown	2.13% (936)	2.26% (774)	1.34% (47)	1.82% (32)	1.82% (34)	2.65% (38)	
Reporting Facility							< 0.0001
Comprehensive	42.72% (18755)	44.24% (15164)	35.07% (1228)	37.21% (1049)	37.01% (692)	43.44% (622)	
Academic	28.06% (12316)	25.88% (8870)	36.29% (1271)	37.21% (1049)	38.02% (711)	28.98% (415)	
Integrated	19.01% (8343)	19.54% (6697)	19.25% (674)	15.22% (429)	16.79% (314)	15.99% (229)	

Patient Characteristics % (N)	Overall 100% (43898)	Non-Hispanic White 78.08% (34275)	Non-Hispanic Black 7.98% (3502)	Hispanic 6.42% (2819)	AAPI 4.26% (1870)	AIAN/ Other/ Unknown 3.26% (1432)	P value
Unknown	3.01% (1322)	2.50% (857)	4.83% (169)	5.21% (147)	4.97% (93)	3.91% (56)	
Charlson/Deyo scores							< 0.0001
0	83.49% (36650)	84.06% (28813)	75.21% (2634)	82.51% (2326)	87.91% (1644)	86.10% (1233)	
1	12.16% (5336)	11.71% (4012)	17.48% (612)	13.16% (371)	10.05% (188)	10.68% (153)	
2–3	4.36% (1912)	4.23% (1450)	7.31% (256)	4.33% (122)	2.03% (38)	3.21% (46)	

Abbreviations: AAPI = Asian American and Pacific Islander; AIAN = American Indian and Alaska Native; NHB = Non-Hispanic Black; NHW = Non-Hispanic White

The median age of the overall population was 61 years (range: 52–68). Of the 43,898 eligible patients included, 26.0% (N = 11422) had a low RS (0–10), 59.9% (N = 26302) intermediate RS (11–25), and 14.1% (N = 6174) had a high RS (> = 26). NHB, Hispanic and AAPI patients were significantly younger at age of diagnosis compared to White patients ($p < 0.001$).

NHB and Hispanic patients were more likely to have household incomes under $38,000 and to rely on Medicaid or be uninsured, whereas NHW and AAPI patients were more often privately insured and had higher household incomes, with over 38% of NHW patients in the ≥$63,000 bracket ($p < 0.001$). Urbanicity patterns revealed that AAPI and Hispanic patients predominantly resided in metropolitan areas, while NHW patients were more likely to live in rural or urban non-metro regions. Regarding facility type, NHB, Hispanic, and AAPI patients were more likely to receive care at academic or comprehensive community cancer programs ($p < 0.001$). Additionally, NHB patients had the highest proportion of Charlson/Deyo comorbidity scores ≥ 1, indicating a greater burden of underlying health conditions ($p < 0.001$, Table 1).

Baseline tumor characteristics are shown in Table 2. Although most patients were node-negative, a higher proportion of NHB and Hispanic patients had lymph node involvement compared to NHW and AAPI patients ($p < 0.001$). NHB and Hispanic patients were also more likely to have tumors greater than 2 centimeters (T2) compared to NHW patients ($p < 0.001$). Histological grade also differed by race/ethnicity with NHB patients exhibiting the highest proportion of grade 3 tumors ($p < 0.0001$).

Median progesterone receptor (PR) expression varied significantly across racial/ethnic subgroups with lower expression in NHB patients ($p < 0.0001$).

Table 2
Baseline Tumor Characteristics

Characteristics % (N)	Overall 100% (43898)	Non-Hispanic White 78.08% (34275)	Non-Hispanic Black 7.98% (3502)	Hispanic 6.42% (2819)	AAPI 4.26% (1870)	AIAN/ Other/ Unknown 3.26% (1432)	P value
Stage							< 0.0001
Stage I	92.37% (40548)	92.85% (31825)	89.38% (3130)	90.35% (2547)	92.19% (1724)	92.32% (1322)	
Stage II	7.05% (3096)	6.65% (2281)	9.22% (323)	9.12% (257)	7.22% (135)	6.98% (100)	
Stage III	0.16% (72)	0.16% (55)	0.26% (9)	0.14% (4)	0.16% (3)	0.07% (1)	
Tumor size							< 0.0001
≤ 1cm	25.07% (11007)	25.91% (8882)	21.10% (739)	21.00% (592)	23.90% (447)	24.23% (347)	
1-3cm	65.38% (28700)	65.00% (22278)	67.65% (2369)	67.22% (1895)	64.97% (1215)	65.85% (943)	
≥ 3cm	9.55% (4191)	9.09% (3115)	11.25% (394)	11.78% (332)	11.12% (208)	9.92% (142)	
Tumor grade							< 0.0001
1	25.91% (11375)	27.13% (9299)	21.30% (746)	21.50% (606)	19.57% (366)	25.00% (358)	
2	56.14% (24646)	56.24% (19275)	51.97% (1820)	57.57% (1623)	58.82% (1100)	57.82% (828)	
3	15.34% (6733)	14.30% (4900)	22.07% (773)	18.09% (510)	18.40% (344)	14.39% (206)	
Hormone status							< 0.0001
ER + Median	95% (IQR 92–99)	95% (93–99)	95% (91–99)	95% (92–100)	95% (95–99)	95% (91–99)	
PR + Median	85% (IQR 35–95)	85% (35–95)	75% (17–95)	85% (40–95)	90% (50–95)	80% (30–95)	

Characteristics % (N)	Overall 100% (43898)	Non-Hispanic White 78.08% (34275)	Non-Hispanic Black 7.98% (3502)	Hispanic 6.42% (2819)	AAPI 4.26% (1870)	AIAN/ Other/ Unknown 3.26% (1432)	P value
Ki-67 status							< 0.0001
Low ≤ 5%	21.82% (9580)	22.39% (7675)	17.50% (613)	20.82% (587)	22.25% (416)	20.18% (289)	
Intermediate 6–29%	59.40% (26075)	59.73% (20471)	56.94% (1994)	58.18% (1640)	58.88% (1101)	60.68% (869)	
High ≥ 30%	18.78% (8243)	17.88% (6129)	25.56% (895)	21.00% (592)	18.88% (353)	19.13% (274)	
Oncotype (RS)							< 0.0001
Low (0–17)	26.02% (11422)	26.49% (9080)	22.27% (780)	24.80% (699)	25.94% (485)	26.40% (378)	
Intermediate (18–25)	59.92% (26302)	60.23% (20643)	56.62% (1983)	60.52% (1706)	59.30% (1109)	60.13% (861)	
High >=26	14.06% (6174)	13.28% (4552)	21.10% (739)	14.69% (414)	14.76% (276)	13.48% (193)	

Abbreviations: AAPI = Asian American and Pacific Islander; AIAN = American Indian and Alaska Native; ER= Estrogen receptor; PR=Progesterone receptor, RS=Recurrence score

Distribution of Ki67 and RS

The overall median 21-gene recurrence score was 15 (IQR: 10–21) with NHB patients having the highest RS at a median of 16 (IQR: 11–24, Fig. 2A). Overall, 34% of NHB patients had a recurrence score >/=26 compared to 22% of NHW patients ($p < 0.0001$). Intermediate Ki67 expression (6–29%) was the most prevalent, comprising the majority of cases in all racial/ethnic groups. The proportion of patients with high Ki67 expression (≥ 30%) was highest among NHB patients and lowest among NHW and AAPI patients ($p < 0.001$). Conversely, NHB patients had the lowest percentage of tumors with low Ki67 expression (≤ 5%) compared to all other racial and ethnic groups ($p < 0.0001$, Fig. 2B).

Agreement between Ki67, RS, and Grade based on Race/Ethnicity

Agreement between RS, Ki67, and grade based on race/ethnicity are presented in Fig. 3. There was slight agreement between Ki67 and RS in the overall population (Kappa = 0.19, $p < 0.0001$), as well as within the low Ki67/low RS (Kappa = 0.07; $p < 0.0001$) and intermediate Ki67/intermediate RS (Kappa = 0.07, $p < 0.0001$) subgroups. In contrast, a fair level of agreement was observed between high Ki67 and high RS (Kappa = 0.351, $p < 0.0001$) in the overall population. When stratified by race/ethnicity, overall agreement between Ki67 and RS remained slight among NHW, Hispanic, and AAPI patients, but was fair among NHB patients (Kappa = 0.2345, $p < 0.0001$). Within the low and intermediate Ki67 and RS categories, agreement was consistently slight across all racial groups ($p < 0.0001$). Among those with high Ki67, fair agreement with high RS was seen across all racial groups, with the highest agreement in NHB patients (Kappa = 0.392), followed by AAPI (Kappa = 0.363), NHW (Kappa = 0.342), and Hispanic patients (Kappa = 0.339), all with $p < 0.0001$.

We also evaluated agreement between RS and grade, as well as Ki67 and grade, in the overall population and racial/ethnic subgroups. Agreement between high RS and high grade was fair in the overall population (Kappa = 0.30, $p < 0.0001$). When stratified by race/ethnicity, agreement was moderate for NHB (Kappa = 0.41, $p < 0.0001$), Hispanic (Kappa = 0.40, $p < 0.0001$), and AAPI (Kappa = 0.38, $p < 0.0001$) patients, but remained fair for NHW patients (Kappa = 0.29, $p < 0.0001$). Agreement between high Ki67 and high grade was moderate overall (Kappa = 0.45, $p < 0.0001$) and remained moderate for NHW (Kappa = 0.44, $p < 0.0001$), NHB (Kappa = 0.46, $p < 0.0001$), and AAPI (Kappa = 0.43, $p < 0.0001$) subgroups, but was lower among Hispanic patients (Kappa = 0.33, $p < 0.0001$, Fig. 3).

Discussion

This large-scale, racially diverse NCDB study of over 43,000 patients with HR-positive, HER2-negative EBC provides critical insights into the relationship between Ki67 expression and RS, and how this relationship may be affected by race/ethnicity. Our study found that NHB patients had significantly higher RS and Ki67 values compared to other racial and ethnic subgroups. NHB patients were also more likely to have higher grade and larger tumors, aligning with previous reports that HR-positive tumors in NHB patients exhibit more aggressive biological features despite similar hormone receptor expression [15–17, 33].

Our findings also concur with other studies that there is only minimal to fair overall concordance between Ki67 and RS [13, 34–35]. Within Ki67 subgroups, concordance with RS was poor in the low and intermediate expression groups, and only fair in the high Ki67 group. This suggests that Ki67 and RS are not interchangeable for risk stratification, except potentially at high levels of tumor proliferation.

Racial stratification revealed that concordance between high Ki67 and high RS was strongest among NHB patients and comparatively lower among NHW and AAPI patients. This may be due to the higher proportion of NHB patients with high Ki67 and high RS values. Nonetheless, these differences suggest that the clinical significance of high proliferation may differ based on ancestry or other biologically relevant factors. Studies indicate that breast tumors in NHB patients, particularly those with HR-positive,

HER2-negative subtypes, exhibit higher genomic instability, increased intratumoral heterogeneity, and distinct mutations such as GATA3 variations compared to tumors in White patients [36–38]. These molecular differences may contribute to more aggressive disease and may partly explain racial disparities in clinical outcomes despite similar Oncotype DX RS.

Multiple socioeconomic and structural determinants may contribute to the higher concordance between high Ki67 and high RS observed among NHB patients. Compared with NHW and AAPI groups, NHB patients in this cohort had lower private insurance coverage, lower neighborhood income and educational attainment, higher comorbidity, and greater care concentration in metropolitan/Southern facilities—factors linked to later presentation and a case mix enriched for higher-grade, more proliferative tumors [39–40]. When RS is obtained primarily for tumors with aggressive clinicopathologic features (e.g., larger size, higher grade), the tested cohort may be enriched for high proliferation [29]. In this context, the NHB subset is more likely to display these features and less uniform testing access, likely explaining the higher probability of concordant high Ki67/high RS [26]. By contrast, broader, routine RS uptake in NHW and AAPI groups may sample more lower-risk and mixed-biology tumors, increasing discordant Ki67/RS pairings and resulting in lower observed concordance.

In our cohort, NHB and Hispanic women were more likely to exhibit high Ki67, larger tumors, and higher-grade disease. These clinicopathologic features represent well-established indicators of aggressive biology and may signal residual risk not fully captured by genomic profiling [41]. This observation is consistent with SEER data demonstrating that NHB and Hispanic women more frequently present with higher-grade, larger, and more proliferative tumors, as well as with NSABP trial analyses showing that traditional pathologic features such as tumor size and grade retain prognostic relevance even when genomic assays are available [23, 26].

Altogether, our results suggest that Ki67 and RS offer distinct but complementary insights into tumor aggressiveness. Integrating both genomic and immunohistochemical markers may enable more accurate risk stratification and guide more equitable treatment recommendations.

The strengths of this study include a large cohort size and inclusion of real-world data from Commission on Cancer–accredited institutions across the United States. However, some limitations must be considered; most prominently, the retrospective design limits the ability to draw causal inferences. Additionally, although the NCDB captures approximately 70% of newly diagnosed cancer cases nationwide, it is restricted to data derived from Commission-accredited hospitals, potentially excluding patients treated in non-accredited or community-based settings. This may result in sampling bias, particularly given the documented underrepresentation of American Indian/Alaska Native, Hispanic, and older populations within the NCDB [42]. These limitations highlight the need for complementary prospective studies with longer follow-up and more inclusive demographic representation to validate and generalize these findings.

Conclusions

In conclusion, this large NCDB analysis demonstrates that Ki67 and the 21-gene recurrence score demonstrate only fair concordance at high expression levels. The relationship between these biomarkers is further modified by race, with non-Hispanic Black patients exhibiting higher RS and Ki67 values and stronger concordance in the high Ki67 and RS subgroup.

Declarations

Funding: *This work was supported in part by grant no.5P30CA196521 from the US National Institutes of Health to support the Tisch Cancer Institute at the Icahn School of Medicine at Mount Sinai.*

Competing Interests: *The authors have no relevant financial or non-financial interests to disclose.*

Author Contributions: *All authors contributed to the study conception and design. Material preparation, data collection and analysis were performed by Deukwoo Kwon, Grace Van Hyfte, Nithya Krishnamurthy, and Rima Patel. The first draft of the manuscript was written by Nithya Krishnamurthy and all authors commented on previous versions of the manuscript. All authors read and approved the final manuscript.*

Data Availability: *The National Cancer Database is available publicly through the American College of Surgeons (https://www.facs.org/quality-programs/cancer/ncdb). The datasets analyzed during the current study can be made available from the corresponding author on reasonable request.*

Ethics approval: *This study was performed in line with the principles of the Declaration of Helsinki. Approval was granted by the Ethics Committee of the Mount Sinai Health System.*

Consent to participate and publish: *Informed consent to participate and for publication of data was obtained from all individual participants included in the National Cancer Database.*

References

1. Dowsett M, Cuzick J, Wale C, Forbes J, Mallon EA, Salter J, Quinn E, Dunbier A, Baum M, Buzdar A, Howell A. Prediction of risk of distant recurrence using the 21-gene recurrence score in node-negative and node-positive postmenopausal patients with breast cancer treated with anastrozole or tamoxifen: a TransATAC study. J Clin Oncol. 2010;28(11):1829–34.
2. Sparano JA, Gray RJ, Makower DF, Pritchard KI, Albain KS, Hayes DF, Geyer CE Jr, Dees EC, Goetz MP, Olson JA Jr, Lively T. Adjuvant chemotherapy guided by a 21-gene expression assay in breast cancer. N Engl J Med. 2018;379(2):111–21.
3. Sparano JA, Gray RJ, Ravdin PM, Makower DF, Pritchard KI, Albain KS, Hayes DF, Geyer CE Jr, Dees EC, Goetz MP, Olson JA Jr. Clinical and genomic risk to guide the use of adjuvant therapy for breast cancer. N Engl J Med. 2019;380(25):2395–405.

4. Rabie MA, Rankin A, Burger A, Youssef MM. The effect of Oncotype DX® on adjuvant chemotherapy treatment decisions in early breast cancer. Annals Royal Coll Surg Engl. 2019;101(8):596–601.
5. Licata L, Viale G, Giuliano M, Curigliano G, Chavez-MacGregor M, Foldi J, Oke O, Collins J, Del Mastro L, Puglisi F, Montemurro F. Oncotype DX results increase concordance in adjuvant chemotherapy recommendations for early-stage breast cancer. NPJ Breast Cancer. 2023;9(1):51.
6. Kalinsky K, Barlow WE, Gralow JR, Meric-Bernstam F, Albain KS, Hayes DF, Lin NU, Perez EA, Goldstein LJ, Chia SK, Dhesy-Thind S. 21-gene assay to inform chemotherapy benefit in node-positive breast cancer. N Engl J Med. 2021;385(25):2336–47.
7. Zhang S, Fitzsimmons KC, Hurvitz SA. Oncotype DX Recurrence Score in premenopausal women. Therapeutic Adv Med Oncol. 2022;14:17588359221081077.
8. Davey MG, Hynes SO, Kerin MJ, Miller N, Lowery AJ. Ki67 as a prognostic biomarker in invasive breast cancer. Cancers. 2021;13(17):4455.
9. Penault-Llorca F, Radosevic-Robin N. Ki67 assessment in breast cancer: an update. Pathology. 2017;49(2):166–71.
10. Nielsen TO, Leung SC, Rimm DL, Dodson A, Acs B, Badve S, Denkert C, Ellis MJ, Fineberg S, Flowers M, Kreipe HH. Assessment of Ki67 in breast cancer: updated recommendations from the international Ki67 in breast cancer working group. JNCI: J Natl Cancer Inst. 2021;113(7):808–19.
11. Harvey J, Thomas C, Wood B, Hardie M, Dessauvagie B, Combrinck M, Frost FA, Sterrett G. Practical issues concerning the implementation of Ki67 proliferative index measurement in breast cancer reporting. Pathology. 2015;47(1):13–20.
12. Leung SCY, Nielsen TO, Zabaglo LA, et al. Analytical validation of a standardised scoring protocol for Ki67 immunohistochemistry on breast cancer excision whole sections: an international multicentre collaboration. Histopathology. 2019;75(2):225–35. 10.1111/his.13880.
13. Patel R, Hovstadius M, Kier MW, Moshier EL, Zimmerman BS, Cascetta K, Jaffer S, Sparano JA, Tiersten A. Correlation of the Ki67 Working Group prognostic risk categories with the Oncotype DX Recurrence Score in early breast cancer. Cancer. 2022;128(20):3602–9.
14. Lee J, Lee YJ, Bae SJ, Baek SH, Kook Y, Cha YJ, Lee JW, Son BH, Ahn SH, Lee HJ, Gong G. Ki67, 21-gene recurrence score, endocrine resistance, and survival in patients with breast cancer. JAMA Netw open. 2023;6(8):e2330961.
15. Rauscher GH, Silva A, Pauls H, Frasor J, Bonini MG, Hoskins K. Racial disparity in survival from estrogen and progesterone receptor-positive breast cancer: implications for reducing breast cancer mortality disparities. Breast Cancer Res Treat. 2017;163:321–30.
16. Lovejoy LA, Shriver CD, Haricharan S, Ellsworth RE. Survival disparities in US Black compared to White women with hormone receptor positive-HER2 negative breast cancer. Int J Environ Res Public Health. 2023;20(4):2903.
17. Sparano JA, Brawley OW. Deconstructing racial and ethnic disparities in breast cancer. JAMA Oncol. 2021;7(3):355–6.

18. Sparano JA, Wang M, Zhao F, Stearns V, Martino S, Ligibel JA, Perez EA, Saphner T, Wolff AC, Sledge GW Jr, Wood WC. Race and hormone receptor–positive breast cancer outcomes in a randomized chemotherapy trial. J Natl Cancer Inst. 2012;104(5):406–14.
19. Makower D, Lin J, Xue X, Sparano JA. Lymphovascular invasion, race, and the 21-gene recurrence score in early estrogen receptor-positive breast cancer. NPJ breast cancer. 2021;7(1):20.
20. Benefield HC, Reeder-Hayes KE, Nichols HB, Calhoun BC, Love MI, Kirk EL, Geradts J, Hoadley KA, Cole SR, Earp HS, Olshan AF. Outcomes of hormone-receptor positive, HER2-negative breast cancers by race and tumor biological features. JNCI cancer Spectr. 2021;5(1):pkaa072.
21. Kim G, Pastoriza JM, Qin J, Lin J, Karagiannis GS, Condeelis JS, Yothers G, Anderson S, Julian T, Entenberg D, Rohan T. Racial disparity in distant recurrence-free survival in patients with localized breast cancer: a pooled analysis of National Surgical Adjuvant Breast and Bowel Project trials. Cancer. 2022;128(14):2728–35.
22. Kantor O, King TA, Freedman RA, Mayer EL, Chavez-MacGregor M, Korde LA, Sparano JA, Mittendorf EA. Racial and ethnic disparities in locoregional recurrence among patients with hormone receptor–positive, node-negative breast cancer: a post hoc analysis of the TAILORx randomized clinical trial. JAMA Surg. 2023;158(6):583–91.
23. Sadigh G, Gray RJ, Sparano JA, Yanez B, Garcia SF, Timsina LR, Obeng-Gyasi S, Gareen I, Sledge GW, Whelan TJ, Cella D. Assessment of racial disparity in survival outcomes for early hormone receptor–positive breast cancer after adjusting for insurance status and neighborhood deprivation: a post hoc analysis of a randomized clinical trial. JAMA Oncol. 2022;8(4):579–86.
24. Kabat GC, Ginsberg M, Sparano JA, Rohan TE. Risk of recurrence and mortality in a multi-ethnic breast cancer population. J Racial Ethnic Health Disparities. 2017;4(6):1181–8.
25. Sparano JA, Zhao F, Martino S, Ligibel JA, Perez EA, Saphner T, Wolff AC, Sledge GW Jr, Wood WC, Davidson NE. Long-term follow-up of the E1199 phase III trial evaluating the role of taxane and schedule in operable breast cancer. J Clin Oncol. 2015;33(21):2353–60.
26. Hoskins KF, Danciu OC, Ko NY, Calip GS. Association of race/ethnicity and the 21-gene recurrence score with breast cancer–specific mortality among US women. JAMA Oncol. 2021;7(3):370–8.
27. Albain KS, Gray RJ, Makower DF, Faghih A, Hayes DF, Geyer CE, Dees EC, Goetz MP, Olson JA, Lively T, Badve SS, Saphner TJ, Wagner LI, Whelan TJ, Ellis MJ, Wood WC, Keane MM, Gomez HL, Reddy PS, Goggins TF, Mayer IA, Brufsky AM, Toppmeyer DL, Kaklamani VG, Berenberg JL, Abrams J, Sledge GW, Sparano JA. Race, ethnicity, and clinical outcomes in hormone receptor–positive, HER2-negative, node-negative breast cancer in the randomized TAILORx trial. J Natl Cancer Inst. 2021;113(4):390–9.
28. Abdou Y, Kantor O, Racz J, Newman L, Pierce LJ, Winer EP. Prognostic and Predictive Insights From Genomic Assays for Breast Cancer in Diverse Populations: A Review. JAMA Oncol. 2025.
29. Guth AA, Chun Kim J, Schwartz S, Montes J, Snyder RA, Axelrod D, Schnabel F. The relationship of race, oncotype DX, and Ki67 in a population highly screened for breast cancer. Breast J. 2017;23(2):177–81.

30. Habermann EB, Day CN, Palis BE, Plichta JK, Wasif N, Weigel RJ, Boughey JC. American College of Surgeons Cancer Programs Annual Report from 2021 Participant User File. Journal of the American College of Surgeons. 2023 Jan 19:10–97.
31. Palis BE, Janczewski LM, Browner AE, Cotler J, Nogueira L, Richardson LC, Benard V, Wilson RJ, Walker N, McCabe RM, Boffa DJ. The national cancer database conforms to the standardized framework for registry and data quality. Ann Surg Oncol. 2024;31(9):5546–59.
32. Shim VC, Baker RJ, Jing W, Puentes R, Agersborg SS, Lee TK, Goreal W, Achacoso N, Lee C, Villasenor M, Lin A. Evaluation of the international Ki67 working group cut point recommendations for early breast cancer: Comparison with 21-gene assay results in a large integrated health care system. Breast Cancer Res Treat. 2024;203(2):281–9.
33. Reid S, Haddad D, Tezak A, Weidner A, Wang X, Mautz B, Moore J, Cadiz S, Zhu Y, Zheng W, Mayer IA. Impact of molecular subtype and race on HR+, HER2 – breast cancer survival. Breast Cancer Res Treat. 2021;189:845–52.
34. Jackisch C, Anastasiadou L, Aulmann S, Argyriadis A, Möbus V, Solbach C, Baier P, Giesecke D, Ackermann S, Schulmeyer E, Gabriel B. The REMAR (Rhein-Main-Registry) real-world study: prospective evaluation of the 21-gene breast recurrence score® assay in addition to Ki67 for adjuvant treatment decisions in early-stage breast cancer. Breast Cancer Res Treat. 2024;207(2):263–74.
35. Crager M, Wijayawardana SR, Gruver AM, Blacklock A, Russell C, Baehner FL, Sapunar F. Population-based estimate for the correlation of the Oncotype DX Breast Recurrence Score® result and Ki67 IHC MIB-1 pharmDx in HR+, HER2–, node-positive early breast cancer. Breast Cancer Res. 2022;24(1):74.
36. Van Alsten SC, Love MI, Calhoun BC, Butler EN, Perou CM, Hoadley KA, Troester MA. Genomic Analysis Reveals Racial and Age-Related Differences in the Somatic Landscape of Breast Cancer and the Association with Socioeconomic Factors. Cancer Res. 2025 Jan 29.
37. Keenan T, Moy B, Mroz EA, Ross K, Niemierko A, Rocco JW, Isakoff S, Ellisen LW, Bardia A. Comparison of the genomic landscape between primary breast cancer in African American versus white women and the association of racial differences with tumor recurrence. J Clin Oncol. 2015;33(31):3621–7.
38. Bacha R, Alwisi N, Ismail R, Pedersen S, Al-Mansoori L. Unveiling GATA3 Signaling Pathways in Health and Disease: Mechanisms, Implications, and Therapeutic Potential. Cells. 2024;13(24):2127.
39. Chen JC, Handley D, Elsaid MI, Fisher JL, Plascak JJ, Anderson L, Tsung C, Beane J, Pawlik TM, Obeng-Gyasi S. Persistent neighborhood poverty and breast cancer outcomes. JAMA Netw Open. 2024;7(8):e2427755.
40. Roy AM, George A, Attwood K, Alaklabi S, Patel A, Omilian AR, Yao S, Gandhi S. Effect of neighborhood deprivation index on breast cancer survival in the United States. Breast Cancer Res Treat. 2023;202(1):139–53.

41. Labidi S, Mulla N, Elkholi IE, Capella MP, Rose AA, Panasci L, Ferrario C, Basik M, Fallah P. High Ki67 expression is associated with increased risk of distant recurrence in Oncotype Dx low risk breast cancer. Clinical Breast Cancer. 2025 Apr 7.
42. Satpathy Y, Nam P, Moldovan M, Murphy JD, Wang L, Derweesh I, Rose BS, Javier-DesLoges J. Comparison of capture rates of the National Cancer Database across race and ethnicity. JAMA Netw Open. 2023;6(12):e2350237.

Figures

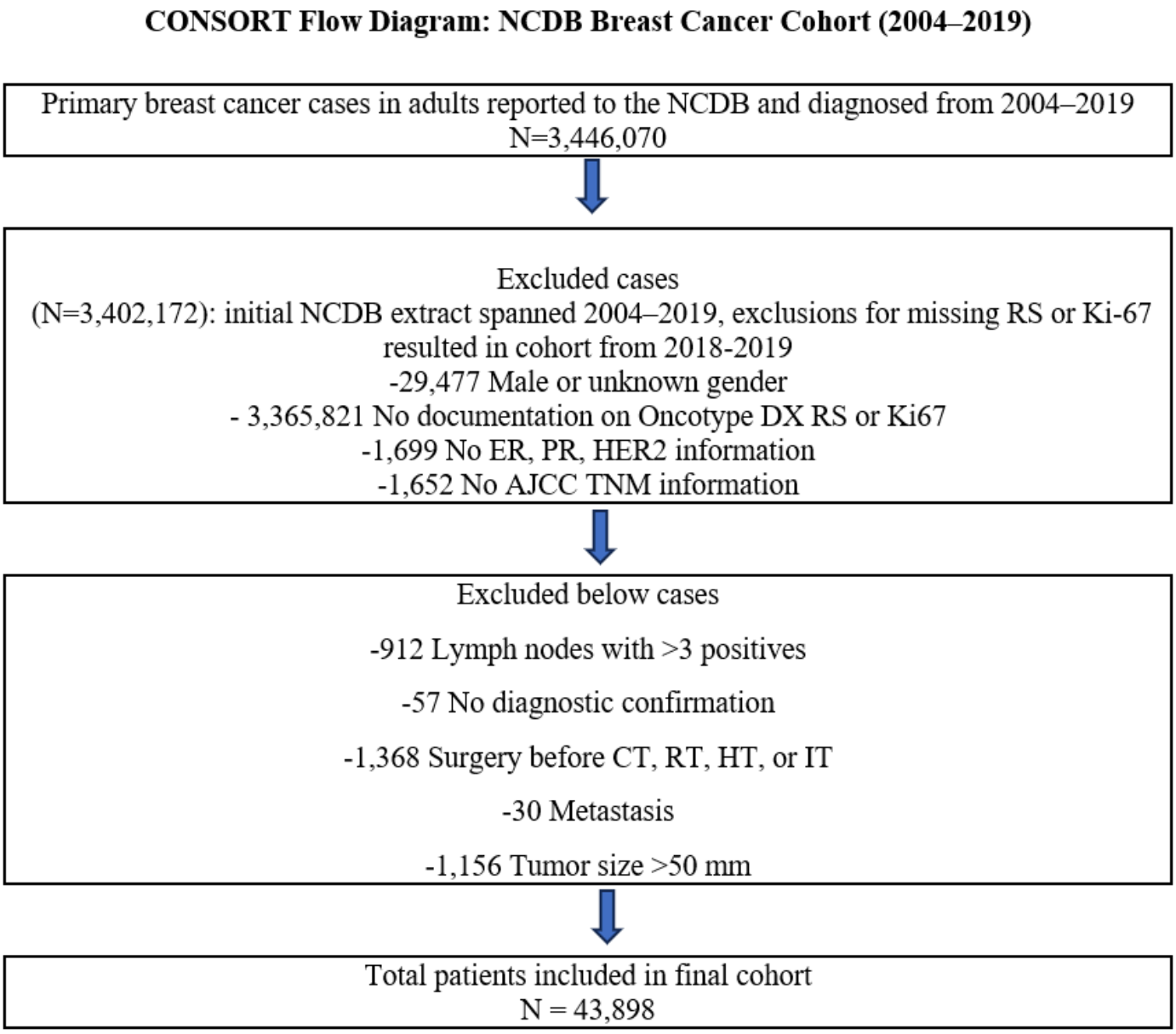


Figure 1

Consort Diagram of Patient Population

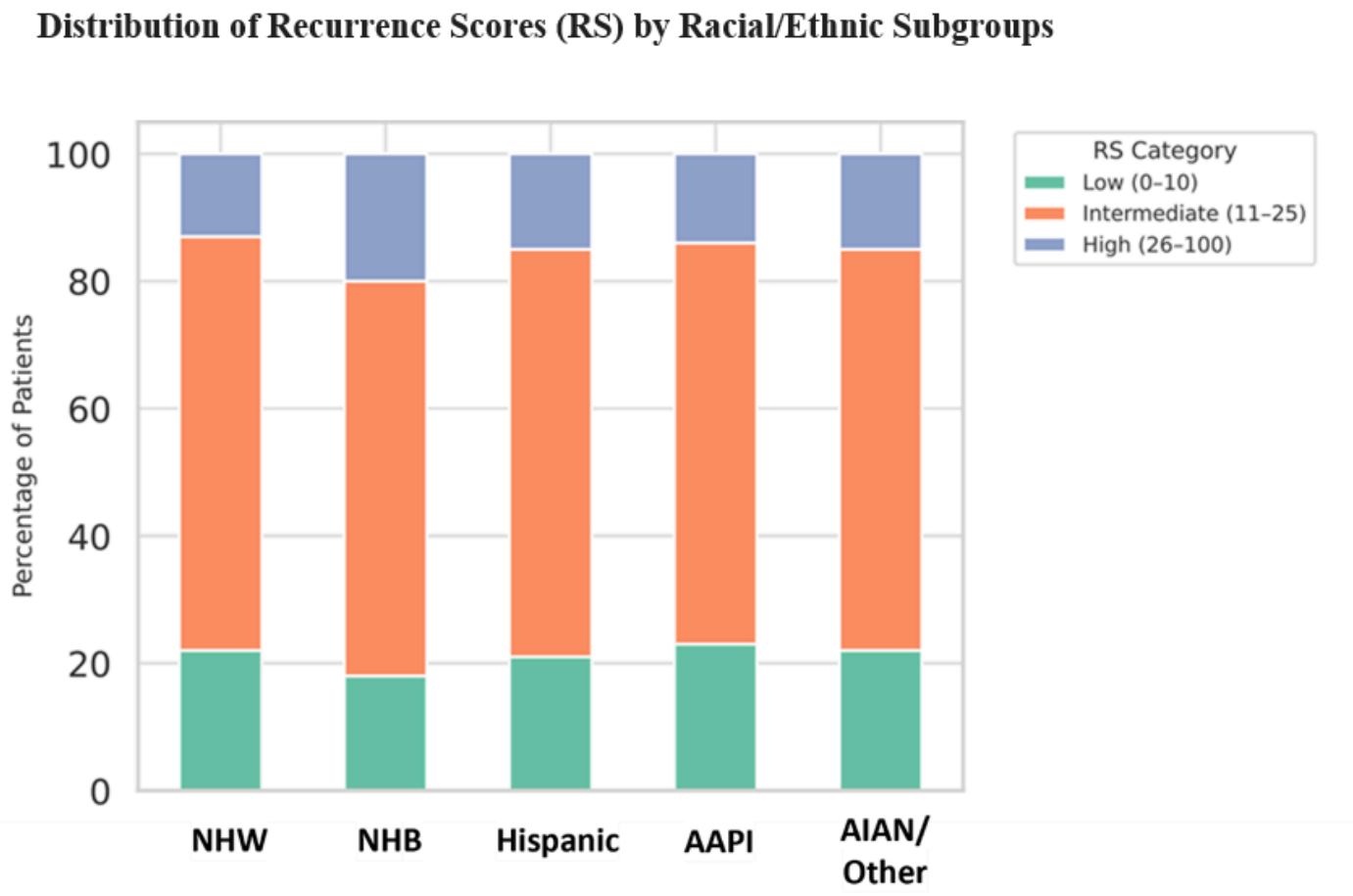


Figure 2

Fig.2A. Distribution of Recurrence Scores (RS) by Racial/Ethnic Subgroups

Across subgroups, most patients had an intermediate RS but a larger proportion of NHB patients had high RS (≥ 26). NHW = Non-Hispanic White, NHB = Non-Hispanic Black, AAPI= Asian American/Pacific Islander; AIAN=American Indian and Alaskan Native.

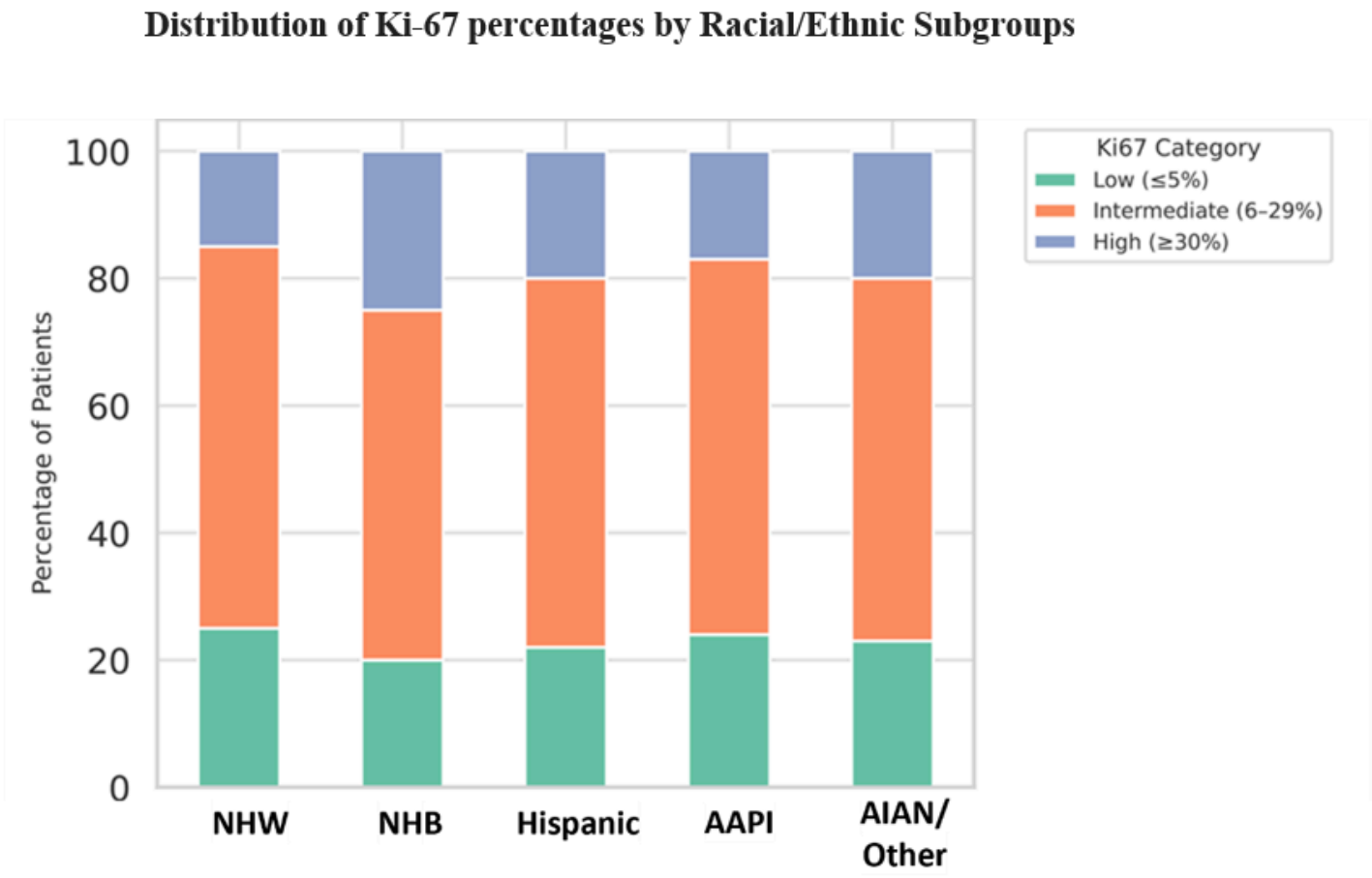


Figure 3

Figure 2B. Distribution of Ki-67 percentages by Racial/Ethnic Subgroups Most patients across racial/ethnic subgroups had an intermediate Ki67 (6-29%) but a larger proportion of NHB patients had high Ki-67% (≥ 30%). NHW = Non-Hispanic White, NHB = Non-Hispanic Black, AAPI= Asian American/Pacific Islander; AIAN=American Indian and Alaskan Native.

Agreement between RS, Ki67, and Grade.

RS and Grade

Risk Category	All	NHW	NHB	Hispanic	AAPI
Overall	0.222	0.209	0.290	0.247	0.246
Low	0.106	0.100	0.140	0.114	0.106
Intermediate	0.067	0.057	0.123	0.107	0.088
High	0.399	0.389	0.439	0.417	0.410

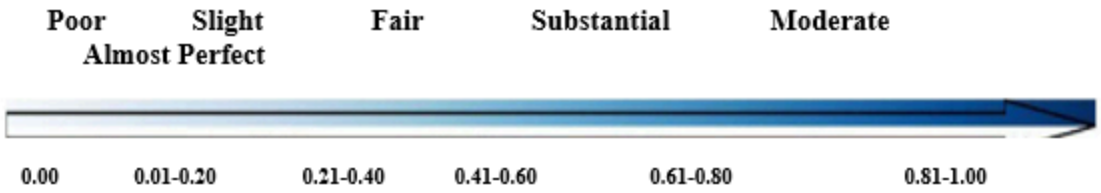
Ki67 and Grade

Risk Category	All	NHW	NHB	Hispanic	AAPI
Overall	0.314	0.307	0.348	0.298	0.353
Low	0.235	0.236	0.221	0.207	0.251
Intermediate	0.155	0.150	0.198	0.134	0.188
High	0.410	0.400	0.452	0.394	0.462

Ki67 and RS

Risk Category	All	NHW	NHB	Hispanic	AAPI
Overall	0.193	0.184	0.235	0.200	0.210
Low	0.069	0.067	0.053†	0.077	0.088
Intermediate	0.066	0.060	0.101	0.077	0.079
High	0.351	0.342	0.392	0.339	0.363

Kappa Interpretation Scale



† $p < 0.0001$ for all Ki67 risk categories except for the low-risk Black population ($p=0.002$)

Figure 4

Figure 3. Agreement between RS, Ki67, and Grade. The figure demonstrates varying levels of agreement based on Kappa coefficients between (A) RS and grade, (B) Ki67 and grade, and (C) Ki67 and RS based on racial/ethnic subgroups. NHW = Non-Hispanic White, NHB = Non-Hispanic Black, AAPI= Asian American/Pacific Islander; AIAN=American Indian and Alaskan Native.