

RESEARCH ARTICLE

<https://doi.org/10.1158/2767-9764.CRC-24-0650>

OPEN ACCESS



Creation and Characterization of a Breast Cancer Tissue Microarray Including Black and White Patients from Florida and Hispanic Patients from Puerto Rico and Florida

Abigail E. Lantz¹, Edna R. Gordián¹, Marilyn Rosa², Marileana Rodríguez-Ruiz³, Joseph O. Johnson⁴, Ryan Gebert¹, Allison Bahr¹, Dung Tsa Chen⁵, Julie Dutil⁶, Jiannong Li⁵, José A. Oliveras Torres⁶, Harold I. Saavedra⁶, Steven A. Eschrich⁵, Idhaliz Flores⁶, and William D. Cress¹

ABSTRACT

Breast cancer is a leading cause of cancer-related mortality among women worldwide and is known to have higher mortality among women with African ancestry. Herein, we describe the creation and characterization of a multiethnic breast cancer tissue microarray (ME-BrTMA) representing tumors from non-Hispanic White ($n = 41$), non-Hispanic Black (NHB; $n = 45$), and Hispanic patients from Puerto Rico ($n = 36$) and Florida ($n = 52$). This ME-BrTMA comprises five blocks with a total of 610 cores: 371 breast cancer tumor cores, 93 breast stromal cores, 96 normal breast tissue cores, 30 non-breast cancer tumor cores, and 20 cores representing normal tissues. Initial characterization of the ME-BrTMA includes standard IHC staining of well-characterized clinical biomarkers, including the estrogen hormone receptors and progesterone hormone receptors, HER2, and Ki-67, interpreted by the coauthoring pathologist (Marilyn Rosa). The IHC results indicated good but imperfect

alignment with clinical diagnoses. Cores from breast cancer tumors from the NHB cohort most frequently scored negative for estrogen receptor (63%, $P < 0.005$) and progesterone receptor (80%, $P < 0.005$) and most frequently have high expression of the Ki-67 proliferation marker (38%, $P < 0.05$). Prediction Analysis of Microarray 50 (PAM50) analysis using RNA from secondary patient blocks showed that the NHB group also most frequently scored in the basal-like category (61%, $P < 0.05$). Taken together, the initial characterization of the ME-BrTMA suggests that it may serve as a representative resource to understand the underlying biology of breast cancer and its relationship to patient outcomes.

Significance: The ME-BrTMA described herein provides a resource that may serve as a tool to understand the underlying biology of breast cancer.

Introduction

The racial and ethnic diversity in cancer research studies including large sequencing efforts (1–6), clinical trials (7–10), and banked tissue resources (11–15) is limited despite evidence that many underrepresented groups,

including Hispanic/Latino (H/L) and non-Hispanic Black (NHB) patients, have poor outcomes relative to non-Hispanic White (NHW) patients (16–18), as well as known differences in genetic drivers (19–22). Cancer disparities are the product of a complex range of socioeconomic, cultural, and biological contributors, of which genomic differences are one factor. The unique biology of many groups has not been adequately characterized because most research resources are built from participants who are NHW with European ancestry (6). A population that is often underrepresented in biological data sources is H/L patients with cancer (2, 4, 6, 7, 12, 23–25). For example, The Cancer Genome Atlas (TCGA), one of the most used cancer databases in research, contains data from approximately 11,000 tumors from 33 cancer types. Yet only approximately 18% of the sequenced tumors, combined, are from underrepresented groups, including only 3% with H/L ethnicity (1). Although ethnicity reporting is improving (10), the number of H/L patients is similarly low with respect to patients with breast cancer. For example, Ding and colleagues (26) noted that only 3.6% of patients with breast cancer (39 of 1,096) reported H/L ethnicity. [ClinicalTrials.gov](https://www.clinicaltrials.gov) did not require the reporting of race/ethnicity information (if even collected) during result submission for trials until 2017 (27). Some previous publications

¹Department of Molecular Oncology, H. Lee Moffitt Cancer Center & Research Institute, Tampa, Florida. ²Department of Pathology, H. Lee Moffitt Cancer Center & Research Institute, Tampa, Florida. ³Medical Doctor Program, Ponce Health Sciences University, Ponce, Puerto Rico. ⁴Analytic Microscopy Core, H. Lee Moffitt Cancer Center & Research Institute, Tampa, Florida. ⁵Department of Biostatistics and Bioinformatics, H. Lee Moffitt Cancer Center & Research Institute, Tampa, Florida. ⁶Department of Basic Sciences, Ponce Research Institute, Ponce Health Sciences University, Ponce, Puerto Rico.

Corresponding Author: William Douglas Cress, H. Lee Moffitt Cancer Center & Research Institute, 12902 Magnolia Drive, Tampa, FL 33612. E-mail: Douglas.Cress@moffitt.org

doi: 10.1158/2767-9764.CRC-24-0650

This open access article is distributed under the Creative Commons Attribution 4.0 International (CC BY 4.0) license.

©2025 The Authors; Published by the American Association for Cancer Research

describing TCGA resources do not even acknowledge H/L or other minority group patients as distinct, instead generalizing them into “other” or “not applicable” categories (28, 29). Even though the TCGA data describe many molecular changes that occur in cancer and have contributed to notable advances in treatment modalities, the data have also widened the gap in understanding molecular traits associated with ancestry-related differences among minority groups (1). Therefore, molecular changes and alterations described so far may not be as relevant in ethnically and racially diverse groups. Additionally, even within the studies that have H/L participants, these patients are often reported as aggregates or homogenous populations, further masking significant differences that can have an impact on their somatic mutation landscape.

As a part of the partnership between Ponce Health Science University in Ponce, Puerto Rico, and Moffitt Cancer Center (MCC) in Tampa, Florida, the Puerto Rico BioBank (PRBB; ref. 14) emerged in 2008 as a biospecimen resource that works to collect, process, and distribute tissues mostly from Puerto Rican individuals, as well as demographic and clinical data for use in cancer research. Its goal is to fill the gap in molecular data needed for translational studies that have an impact on H/L women, specifically the population of Hispanics in Puerto Rico (HPR). Most studies indicate that H/L and NHB women with breast cancer have significantly poorer survival outcomes (30–35) relative to NHW women. However, these findings are not universal, and some studies find that certain Hispanic groups have better prognostic outcomes than NHW women with breast cancer (31–52).

In Puerto Rico, breast cancer has been the most frequently diagnosed neoplasia in women and the cancer type with the highest mortality (50). Additionally, patients with breast cancer in Puerto Rico experience a worse 5-year survival rate after diagnosis that, so far, has not been directly associated with IHC surrogate markers, when compared with NHWs (52). After multivariate adjustments, Ooi and colleagues (31) showed that Puerto Ricans have significantly higher HRs than other H/Ls despite being given appropriate treatment. Puerto Rican patients with breast cancer in the United States are usually grouped in a broad Hispanic cohort, along with women from other very diverse H/L origins [including other Caribbean Hispanics (Dominican and Cuban), Central American Hispanics (including Mexican), and South American Hispanics and Latinos], generalizing the group as homogenous. However, some differences in risk factors among Puerto Rican patients with breast cancer in the United States, other Hispanic ethnicities, and Puerto Rican patients with breast cancer in Puerto Rico have been noted, warranting further investigations (51).

In previous work (53), we leveraged cancer registry and patient questionnaire data to describe four cohorts of patients with breast cancer from Florida and Puerto Rico. Herein, we use available tissue blocks representing a subset of patients from the previously described cohort (53) to construct a multiethnic breast cancer tissue microarray (ME-BrTMA). The ME-BrTMA contains samples from NHW, NHB, and Hispanic patients in Florida (HF), as well as from HPR patients. Supplementary Tables S1 and S2, within the Supplementary Information summarize the clinical and outcome characteristics of each cohort. Figure 1 highlights the construction of the ME-BrTMA (Fig. 1A) and how it fits into our larger goal to characterize (Fig. 1B) four cohorts of patients with breast cancer (Fig. 1C) at five levels (Fig. 1D), including cancer registry data (53), patient questionnaire data (53), clinical data, molecular data (in progress), and at the level of a representative TMA,

allowing for spatial proteomic analysis using multiplex immunofluorescence. It is our intention that this TMA will support hypothesis-driven biomolecular research to understand breast cancer outcomes while representing multiple racial and ethnic groups.

Materials and Methods

Ethics statement

This is to certify that the Institutional Review Board (IRB) has reviewed and approved the research protocol titled “The Puerto Rico Biobank-PRBB” (IRB#00033379, Advarra IRB#00000971). The study has been evaluated for its ethical considerations, potential risks, and benefits to participants. The approval is granted based on the adherence to ethical guidelines and regulatory requirements of the U.S. Common Rule.

The TMA

Tissue samples in the form of surgical blocks were collected from female patients with primary breast cancer diagnosed between 1998 and 2019, as well as from non-breast tumor controls and non-breast cancer healthy controls. Patients were selected from four race and ethnicity cohorts, defined as follows: NHW patients were consented to the Total Cancer Care (TCC) at the MCC, donated a tissue block with tumor, and identified their ethnicity as “non-Spanish” and their race as “White.” NHB patients consented to the TCC at the MCC in Florida, donated a tissue block with tumor, and identified their ethnicity as “non-Spanish” and their race as “Black.” HF patients consented to the TCC at the MCC, donated a tissue block with tumor, and identified their ethnicity as a specified or unspecified Spanish origin. HPR patients consented to the PRBB in Puerto Rico and donated a tissue block with tumor. Some patient data were acquired from the Puerto Rico Central Cancer Registry (RRID: SCR_023507). It is important to note that the H/L patients in our cohorts may be of any race. Of 52 patients in the HF group, 42 self-reported to be of the White race. In the HPR group, 22 patients of 36 self-reported to be of the White race. The patients did not self-report being on any other races. All tissue samples were collected after written informed consent under IRB-approved banking protocols based on adherence to ethical guidelines and regulatory requirements of the U.S. Common Rule. The collection of the HPR cohort by the PRBB was covered by IRB#00033379 and IRB#00000971. NHW, NHB, and HF patient samples were acquired under the TCC Protocol (NCT03977402; Advarra IRB# Pro00014441).

Because their numbers were limiting, blocks from NHB and HPR patients were identified first and all were used. Blocks from White patients, available in excess, were then matched to NHB and HF patients by age at diagnosis, body mass index, and pathologic T stage to the extent that the available data and tissue samples allowed. Among matched NHW patients, those with the most available tumor tissue were selected. An experienced breast cancer pathologist used fresh hematoxylin and eosin slides to mark the blocks for punching 1.0-mm cores of stroma, normal tissue, and two or more tumor cores. All tumor punches were performed from areas of greater than >30% cellularity. Treatment status was unknown prior to block selection. As was possible, two tumor cores, one stroma core, and one normal core were acquired from each patient, though there was variation in the amount and type of available tissue, and 81 patients are represented by three or more tumor cores. The ME-BrTMA was assembled in five individual blocks, with

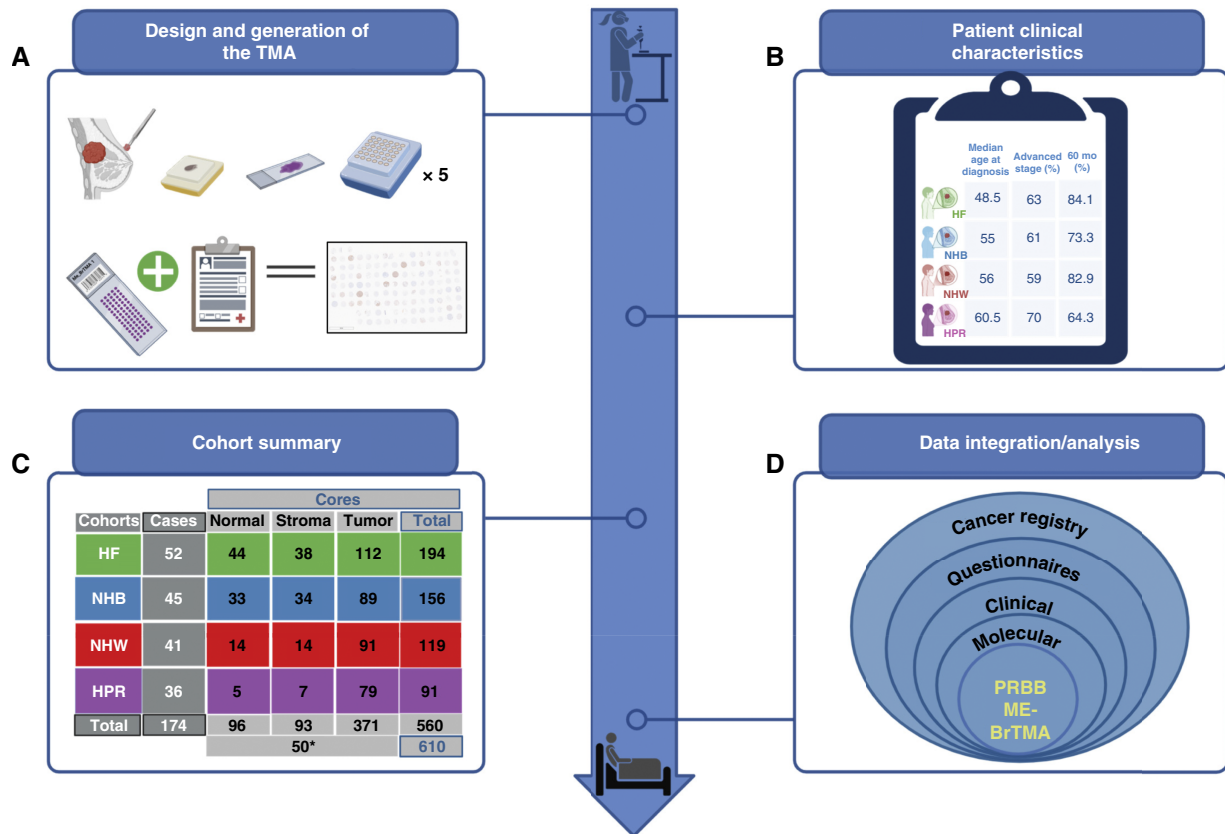


FIGURE 1 The multiethnic breast cancer cohort study. This study represents a long-term commitment of the PRBB to address factors that drive poor breast cancer outcomes within understudied populations. **A**, The ME-BrTMA was constructed from available formalin-fixed, paraffin-embedded breast tissues of well-annotated, consented study participants from the MCC-TCC and the PRBB. **B**, General patient clinical characteristics of the four cohorts (more details available in Supplementary Table S1). The TMA represents four groups: HF, HPR, NHB, and NHW. **C**, The final cohorts as represented by cores on the TMA include 45 NHB, 52 HF, and 41 NHW patients from the MCC and 36 HPR patients from the PRBB. The TMA is composed of five blocks with 610 total cores: 560 total cores from breast tissue (96 normal, 371 tumor, and 93 stromal cores) and 10 control cases on each block (50*). See Supplementary Table S3 for details. **D**, A representation of the levels of data that we seek to obtain as part of the broader study. Created in BioRender.com. Gordian, E. (2025) <https://BioRender.com/l83r702> with additional edits in PowerPoint.

135 to 150 tissue cores per block in the MCC Tissue Core Facility (RRID: SCR_012364). Randomization was performed among 610 cores across five blocks to avoid batch effects. Of the tissue cores included within each block, six are non-breast tumor controls (ovary, colon, pancreas, prostate, kidney, and brain) and four are non-breast normal tissue controls (kidney, salivary gland, colon, and tonsil). See Supplementary Table S3 for an overall summary of the 610 tissue cores of the TMA.

Staining and analysis of the TMA

The ME-BrTMA was stained for hematoxylin and eosin, estrogen receptor (ER), progesterone receptor (PR), HER2, and Ki-67. The specific antibodies used were as follows: (i) ER (Ventana, cat. # 790-4324, RRID: AB_2857956), (ii) PR (Ventana, cat. # 790-2223, RRID: AB_2335976), (iii) Ki-67 (Ventana, cat. # 790-4286, RRID: AB_2631262), and HER2 (Cell Signaling Technology, cat. #4290, RRID: AB_10557104). Additional technical details are included in Supplementary Table S4, and detailed results for the 371 tumor cores are available in Supplementary Table S5. Scoring for ER/PR and HER2 was

interpreted using the current clinical guidelines of the American Society of Clinical Oncology and College of American Pathologists. The five blocks of the TMA were scored individually by a clinical pathologist (author Marilyn Rosa) who was blinded to the cohort (race/ethnicity grouping) identification. The ER was considered negative with less than 1% positively staining nuclei, low-positive with 1% to 10% positively staining nuclei, and positive with greater than 10% positively staining nuclei. For PR, samples with less than 1% positively staining nuclei were considered negative, and those with 1% or greater positively staining nuclei were considered positive. HER2 0 and 1+ were considered negative, HER2 2+ was considered equivocal, and HER2 3+ was considered positive (54, 55). For HR status, only cases with both ER and PR status were considered; HR– were ER– and PR– cases, and all others (ER–low-positive or ER–positive with a PR–positive result) were HR+. Clinical ER, PR, and HER2 results were acquired from cancer registry data sources to ensure that the tissue samples within the ME-BrTMA adequately represented the corresponding patient's disease. Clinical HR was determined with only cases with both clinical ER and clinical PR status. To quantify the frequency of HR/HER2 subtypes, ER and PR were considered negative if the

percentage of nuclear staining was less than 1% and positive with 1% or greater positive nuclei. HER2 was considered negative with a score of 0 or 1+, equivocal with a score of 2+, and positive with a score of 3+. Clinical HR/HER2 subtypes were determined with HER2 borderline or positive considered as HER2+.

Statistical analysis

Statistical analysis of staining results was conducted using descriptive statistics, Fisher exact test, χ^2 test, and Kruskal–Wallis test, as appropriate. Analysis was performed using R statistical software (version 4.2.0, R Core Team 2022; R Project for Statistical Computing, RRID: SCR_001905). χ^2 tests were used to compare proportions, unless expected counts for categories were <5, in which case the Fisher exact test is used. All statistical testing and visualization code are publicly available at <https://github.com/Moffitt-Cancer-Center/me-brtma-summary> (GitHub, RRID: SCR_002630).

Comparison of ER, HR/HER2, and PAM50 subtypes were performed on categorical data by factor level using the Fisher exact test in R version 4.2.0. Comparisons of PR, HR, and HER2 were performed on categorical data using the χ^2 test. Analysis of Ki-67 positivity (represented as continuous percentages) was compared among cohorts (NHW, NHB, HF, and HPR) using the Kruskal–Wallis H test in R version 4.2.0. No *post hoc* tests were used for subsequent evaluation.

Gene expression analysis

Tumor RNA was acquired from patients with sufficient tissue samples for RNA extraction. RNA extraction was performed from a second formalin-fixed, paraffin-embedded block, where available. All RNA samples were DNase-treated and eluted in water. The samples were processed using Illumina TruSeq RNA Exome Kit and sequenced on one NovaSeq-600 S2-200 sequencing run (101 × 101). PAM50 gene expression signatures were calculated using the *genefu* package in R (56), and patients were classified according to their tumor-intrinsic molecular subtype into the following categories: basal-like, HER2-enriched, luminal A, luminal B, and normal-like.

Data availability

All data are available within the main article and supplementary files. The data and analysis generated in this study and select TMA core images are available within this article's Supplementary Tables S1–S15. Additional quality control analyses are present in Supplementary Tables S16–S21. Individual TMA and core images are available upon request from the corresponding author.

Results

The frequency of ER-negative and PR-negative expression is highest in tumor cores from patients from the NHB cohort

The five blocks of the TMA were sectioned and stained with clinically relevant antibodies against ER, PR, HER2, and Ki-67 and were scored individually by a clinical pathologist (author Marilyn Rosa). Supplementary Figures S1 through S4 provide examples of positively and negatively stained cores for each of the four antibodies from each cohort. To provide an initial characterization of the four cohorts represented on the ME-BrTMA, we

compared the ER (Fig. 2A) and PR (Fig. 2B) expression among the cores of the four cohorts (Supplementary Table S6). The NHB cohort had the highest fraction of patients with ER-negative cores (63%, overall $P = 0.001$, Fisher exact test), PR-negative cores (80%, overall $P = 0.003$), and combined HR-negative (59%, overall $P = 0.006$) cores.

Using the determined ER and PR scores from individual cores, we sought to determine how well the values aligned with the patient-level clinical results. Supplementary Table S7 summarizes the results for the ER. For this analysis, we simplified the ER classification to positive or negative as the classification of ER stains as negative, low-positive, and positive is a new development for diagnostic purposes. Our existing clinical classifications do not include this level of resolution; therefore, we can compare the simplified ER stain results with the clinical ER status. Note that we also remove the equivocal from the clinical results. Of the 135 successfully evaluated cores for which we had documented ER-negative clinical results, 133 matched the expectations from the clinical reports. In contrast, two cores were low-positive. Of the 173 cores from ER-positive patients, 19 (11%) were ER-negative. Supplementary Table S8 summarizes the results of the PR cohort. Of the 165 cores from PR-negative patients, five cores (representing four patients) were scored positive. Of the 145 cores from PR-positive patients, 49 (34%) were scored negative in the TMA. Figure 2C and Supplementary Table S9 summarize the results for the combined ER/PR (HR). Of the 126 cores from HR-negative patients, four (3%) were positive, with ER and/or PR expressing 1% or greater positive nuclei. Of the 178 cores from clinically HR-positive patients, 25 (14%) were negative.

Triple-negative (HR–/HER2–) is more frequently found in the NHB cohort

The expression of the HER2 receptor is a critical predictive and prognostic clinical biomarker in breast cancer, and thus we stained our TMA with an HER2 antibody to estimate HER2 expression using published criteria (54). Supplementary Figure S3 highlights examples of HER2-stained cores from each cohort that were scored as HER2 0, HER2 1+, HER2 2+, or HER2 3+. Figure 3A shows that the frequency of HER2+ staining is similar by cohort with no statistically significant differences noted ($P = 0.387$, χ^2 test). Supplementary Table S10 shows that of the 182 cores from HER2-negative patients, all were scored as either negative ($n = 103$, 57%) or equivocal ($n = 79$, 43%). Among 35 cores from HER2+ patients, 20 (57%) were scored as HER2 3+, and 15 were scored as negative ($n = 8$, 23%) or equivocal ($n = 7$, 20%). Patients with negative (HER2 0+) and positive (HER2 3+) staining have a straightforward therapeutic implication when compared with equivocal or the most recent clinical subclassification of HER2-low (IHC 1+ or IHC 2+ *in situ* hybridization nonamplified; ref. 57). Disagreements with whole slides stained for clinical purposes and core-based scoring are expected because the core size is significantly smaller and, therefore, less representative than the complete pathologic samples.

Figure 3B highlights our findings after aggregating ER, PR, and HER2 scores into distinct HR/HER2 subtypes; a significant difference in the frequency of each subtype by cohort was noted ($P = 0.006$, Fisher exact test, see Supplementary Table S11). The aggressive HR–/HER2– (triple-negative) disease comprised more than half (56%) of breast tumor cores from the NHB cohort, demonstrating a higher frequency of this

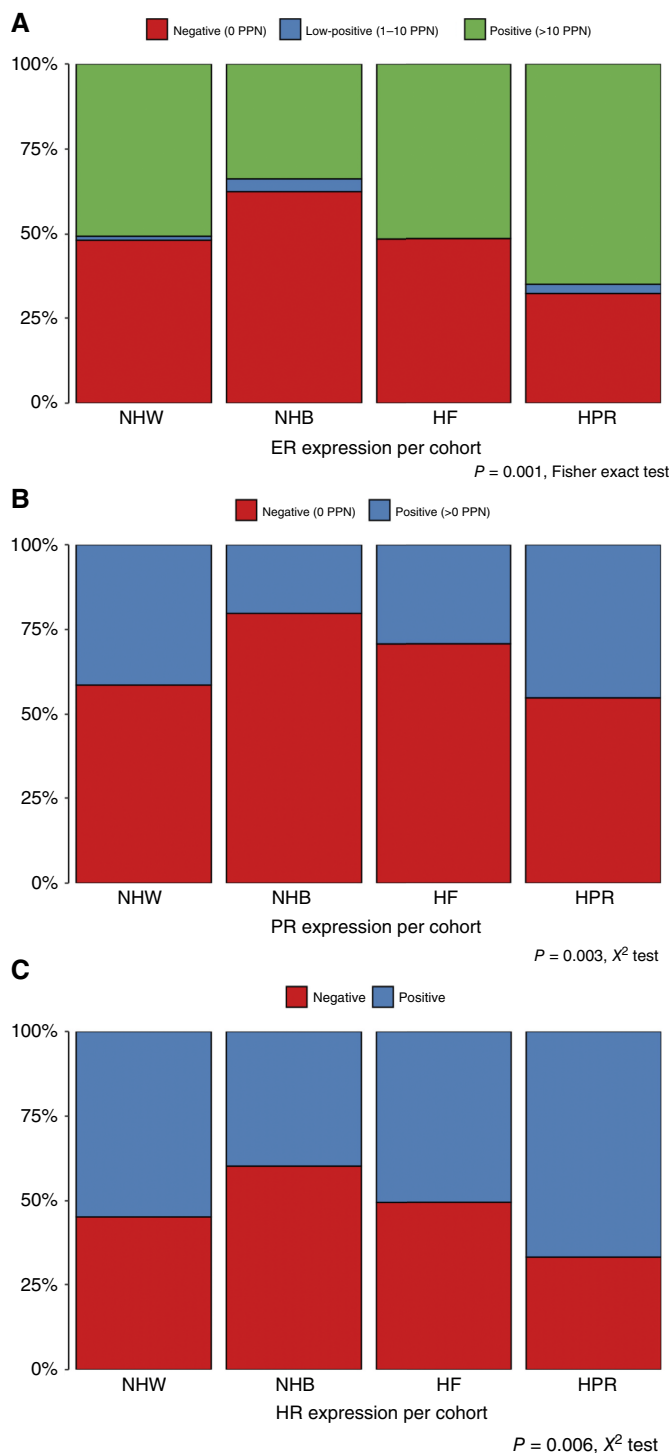


FIGURE 2 The frequency of ER-negative and PR-negative expression is highest in tumor cores from patients in the NHB cohort. **A**, Stacked bar chart showing that the frequency of negative ER expression is highest in cores from the NHB cohort (Fisher exact test, $P = 0.001$). **B**, Stacked bar chart showing that the frequency of PR-negative expression is highest in cores from the NHB cohort (χ^2 test, $P = 0.003$). **C**, Stacked bar chart showing that the frequency of HR-negative expression is highest in cores from the NHB cohort (χ^2 test, $P = 0.006$, χ^2 test). PPN, percent positive nuclei.

subtype than in any other cohort (Fig. 3B). The HR+/HER2− subtype, which tends to be less aggressive and more readily treatable, was most frequently found among HPR patients (64%). Of the 80 cores associated with HR−/HER2− clinical results, 77 were also HR−/HER2− on pathology reading (Supplementary Table S12).

High expression of the proliferative marker Ki-67 is more frequent among cores from the NHB cohort

Ki-67 is a strong proliferative marker in breast cancer and can distinguish between the luminal A and the more aggressive luminal B subtype. Supplementary Figure S4 shows examples of Ki-67–stained cores from each cohort by high and low staining. For Ki-67, we chose a high-to-low risk cutoff of $\geq 14\%$ Ki-67+ nuclei, based on previous reports of its demonstrated prognostic significance (58, 59). There was a significant difference ($P = 0.033$, Kruskal–Wallis test, statistic = 8.74, 3 degrees of freedom) in Ki-67 expression observed among the four cohorts. Fig. 4A shows the increased frequency of high expression of this proliferative marker among the NHB women in our study relative to the NHW, HF, and HPR women. Figure 4B demonstrates the fraction of patients in each cohort above and below the 14 percent positive nuclei cutoff. Supplementary Figure S5 and Supplementary Table S13 extend the analysis of Fig. 4, summarizing Ki-67 percent positivity by clinical subtype and by cohort.

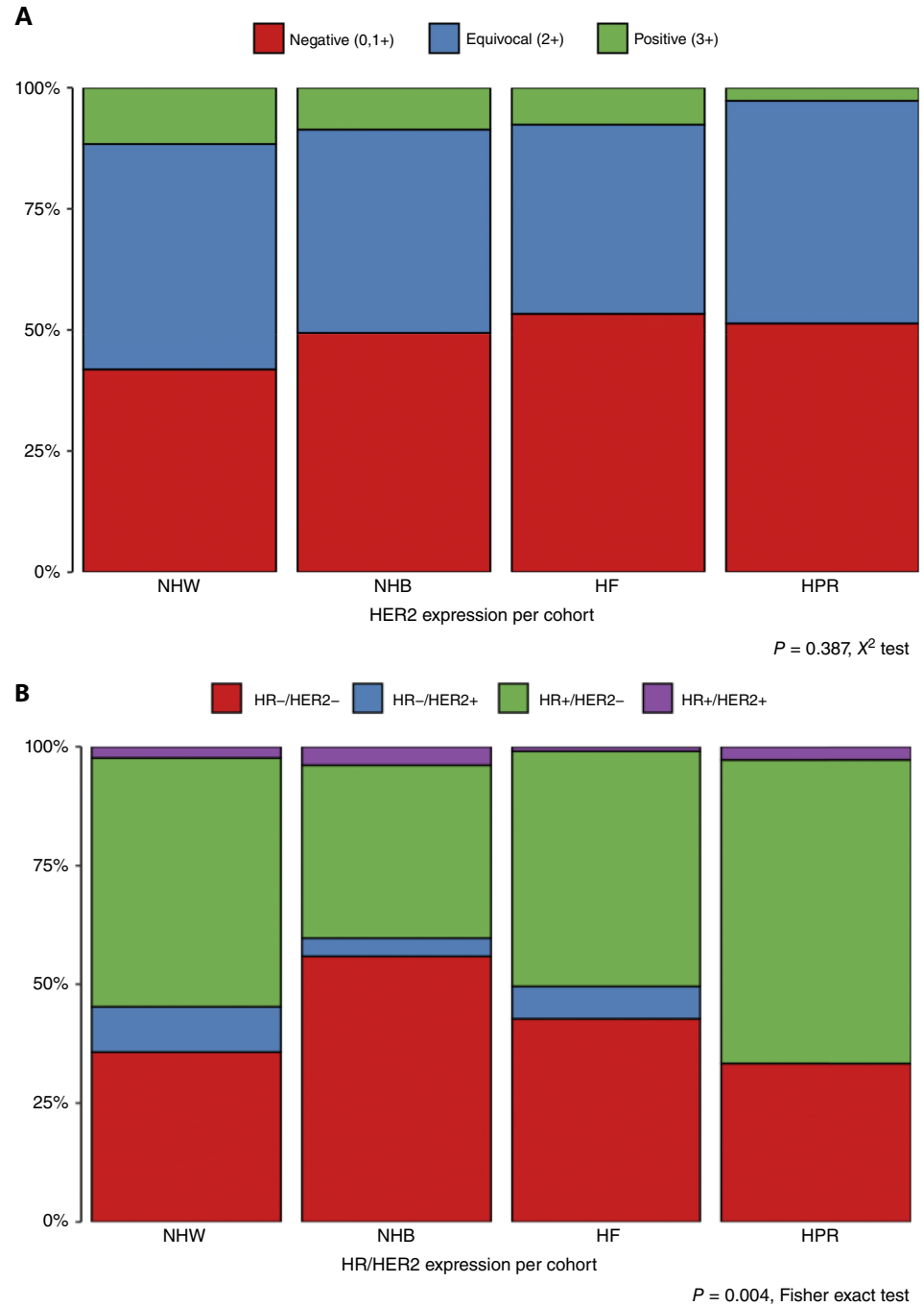
PAM50 analysis shows that the basal-like subtype is more frequent in the NHB patient cohort compared with the Hispanic cohorts

On comparison of the PAM50 signature results by cohort, a significant difference ($P = 0.031$, Fisher exact test) in subtype frequency was observed. These data represent a subset of ME-BrTMA patients with adequate duplicate blocks available to us for RNA extraction and do not include any women from the NHW cohort. The aggressive basal-like subtype is more frequently found among NHB patients (61%) than among those from the HF (34%) and HPR (20%) cohorts. The 61% basal-like breast cancer observed in NHB women is close to the 56% triple-negative breast cancers (TNBC) observed in this cohort (Fig. 5; Supplementary Table S14). Though basal-like breast cancers and TNBCs share many features, not all basal-like subtypes are TNBCs (60). Therefore, the ~5% difference observed in this cohort could represent additional molecular traits that can further help elucidate its aggressive phenotype. Luminal A subtype, which has a better prognosis than other subtypes, is more frequently found in the HPR cohort (60%) than in the HF (31%) and NHB (17%) cohorts. Supplementary Table S15 provides a summary of PAM50 subtypes by HR/HER2 status and cohort.

Additional quality control analysis demonstrates that the tissue cores of the ME-BrTMA represent the patients included

Two additional quality control analyses were performed to demonstrate that the ME-BrTMA is representative of the patients included, as recommended by the reviewers of this article. Supplementary Tables S16–S21 examined ER, PR, and HER2, respectively. In the first approach, we examined how well high and low individual tumor scores matched among all patients. In such tabulations, the largest values fall along the diagonal, indicating consistency

FIGURE 3 TNBC is most common in the NHB cohort. **A**, Stacked bar chart showing the fraction of patients with each level of HER2 expression among the four cohorts ($P = 0.387$, χ^2 test). **B**, Stacked bar chart showing that the HR-/HER2- (triple-negative) phenotype is more frequent in the cores from the NHB patients ($P = 0.004$, Fisher exact test). HR, hormone receptor (ER and PR); PPN, percent positive nuclei.



in scoring. Supplementary Tables S16, S18, and S20 reveal that in every case, there is excellent consistency. However, as expected, when 1-mm cores are used to represent an entire tumor, there are discrepancies. ER consistency was the best with only 5 discrepancies, PR had 15 discrepancies, and HER2 had 30 discrepancies. In the second approach, we tabulated how many cores represent each patient, broken down by cohort. Supplementary Tables S17, S19, and S21 reveal that 45% of the patients are represented by three or more tumor cores, 30% are represented by two cores, 24% are represented by one core, and 2.4% are not represented by tumor core because of manufacturing errors.

Discussion

Through future studies performed by our colleagues and investigators in the community, the ME-BrTMA represents a tool for diversifying translational breast cancer research. Previous TMAs (13, 61) have allowed us to detect levels of mitotic kinases in breast cancer tissues of patients of different races and ethnicities and correlate their levels with pathologic subtypes of breast cancer (e.g., PR, ER, HER2, and Ki-67), as well as surrogate markers of metastasis (e.g., vimentin). In addition to the utility inherent to a TMA, the ME-BrTMA is noteworthy because of the racial and ethnic diversity of

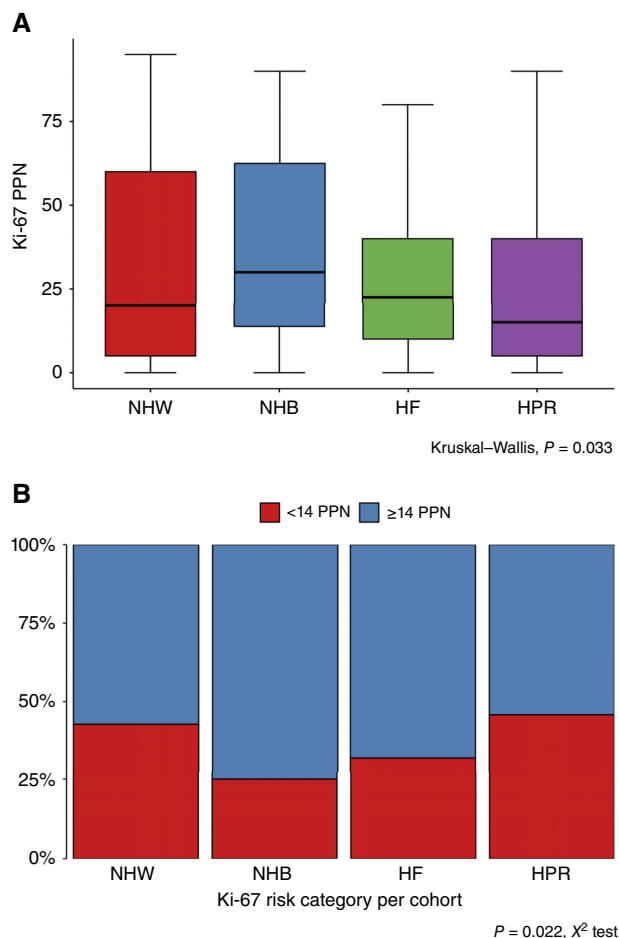


FIGURE 4 High expression of the proliferative marker Ki-67 is more frequent among cores from the NHB cohort. **A**, Boxplots comparing Ki-67% positivity as a function of the cohort. **B**, Risk categories for Ki-67 (low risk, <14 PPN vs. high risk >14 PPN) are shown as stacked bar charts. Supplementary Table S13 tabulates these data (see rows Ki-67 PPN and Ki-67 risk category). PPN, percent positive nuclei.

patients from whom tissue was acquired, the robust patient-level data available, and its representation of known biological differences in the four groups.

The expression of the ER and PR hormone receptors is a critical prognostic and predictive clinical biomarker in breast cancer. Patients with ER/PR+ breast cancers are candidates for endocrine therapy with tamoxifen as adjuvant therapy. The National Comprehensive Cancer Network guidelines currently recommend tamoxifen or aromatase inhibitors for postmenopausal patients with breast cancer as an option to prevent recurrence and improve survival. However, clinical trials have shown that breast cancer recurrence after 5 years of adjuvant endocrine therapy can range from 10% to 41% (32), suggesting the need to further stratify patient responses, considering their unique tumor biology along with their genetic ancestry. In addition to ER and PR, HER2 testing also guides treatment options for patients with breast cancer. Several decades back, HER2+ tumors were considered aggressive as TNBC. However, the discovery of mAb treatments, such as trastuzumab, in the early 1980s changed the outcome of early diagnosed patients and patients that exhibit HER2+, *de novo* metastatic disease (62).

We have previously analyzed HR/HER2 expression in broader cohorts of NHW, NHB, HF, and HPR patients and found the prevalence of HR+/HER2+ disease in our sample to reflect the frequency at which this subtype has been found in the literature (53). However, the patients included within the ME-BrTMA represent a sample of tumors with a higher frequency of HR+/HER2+ than we may have expected (33%–56%), given that only 12% to 27% of patients in each race and ethnicity group in our previous study were HR+/HER2+. Additionally, although we previously found a significantly lower frequency of HR+/HER2+ disease in HPR patients relative to NHW patients, this subtype was most frequently found among HPR patients within the ME-BrTMA compared with each of the other three cohorts. We expect that these differences in frequency relate to the fact that patients with aggressive disease are more likely to seek treatment at academic hospitals and are more likely to consent to research protocols than patients who are represented in cancer registries.

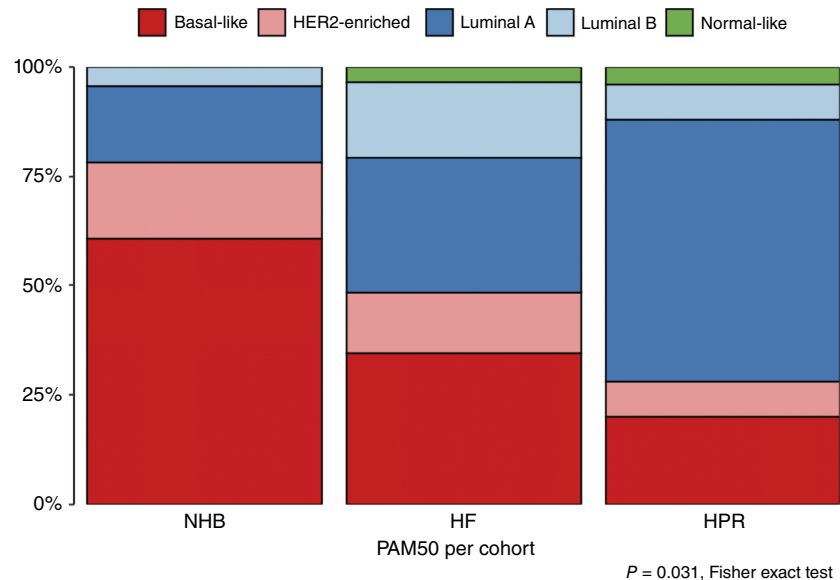
Clinical, pathologic, and molecular data obtained from investigations using the ME-BrTMA will also provide information on another crucial factor affecting these groups: treatment response. It is important to highlight that limited understanding is available about the susceptibility to chemotherapy agents among certain populations. Previous reports have addressed the importance of finding ethnicity-specific genes or gene signatures that can explain response, resistance, or toxicities to cancer compounds such as alkylating agents or mAbs (trastuzumab) used for treating breast cancer in HER2+ patients (8, 9).

As a measure of proliferation, Ki-67 expression can be used for breast cancer prognosis and to predict response to chemotherapeutic drugs. This marker is especially useful in the setting of HR+/HER2+ disease because it can be used to differentiate between luminal subtypes (63). However, its use has not been completely adopted as standard of care in pathologic reporting because of the inconsistencies in surgical acquisition, IHC analysis, and cutoff values. Having said that, Ki-67 is considered an informative marker that can aid in prognosis as well as clinical decisions if shown to improve these in conjunction with other molecular markers (64). The degree of Ki-67 expression has been seen to vary by racial/ethnic group, with significantly higher Ki-67 scores for NHB patients relative to NHW patients (65).

To begin focusing on the unique biological features of each minority group, the PAM50 panel results were analyzed for a subset of the NHB, HF, and HPR groups. NHB and HPR patients show an inverse expression of the most and least aggressive subtypes, with most tumors from NHB patients being of the aggressive basal-like subtype and most tumors from the HPR patients being of the more favorable luminal A subtype. Previously, we noted a similarly poor 5-year overall survival in populations of NHB and HPR patients (53). In combination with this marked difference in PAM50 subtypes between NHB and HPR patients in the ME-BrTMA, these findings warrant further investigation into the biological and social drivers of poor survival among HPR patients.

As additional data resources become available, we aim to further characterize the patients from whom ME-BrTMA tissues were acquired. Genetic ancestry analysis will allow us to better represent the relationship between various biological markers and a patient's heritage, without relying on reported race and ethnicity as proxies for true ancestry. We will also have a broader array of clinical data, particularly about treatment regimens and timing, because these contribute to patient prognosis and, in neoadjuvant chemotherapy, can affect the ME-BrTMA tissue samples. Continuous improvements are being made in data sharing and accessibility to encourage further utilization of this racially and ethnically diverse resource.

FIGURE 5 PAM50 analysis shows that basal-like subtype is most frequent among NHBs. Stacked bar chart shows the results of PAM50 gene expression analysis on patients represented on the ME-BrTMA with available RNA sequencing data. The data show that the basal-like subtype is most frequent among NHBs represented on the TMA ($P = 0.031$). Supplementary Table S15 tabulates the PAM50 subgroup breakdown by cohort. Note that RNA sequencing data for the NHW patients in cohort 1 are not included because they are not available.



In conclusion, the unique breast tumor biology of Puerto Rican patients represents a significant gap in the breast cancer literature. We have created a diverse, well-annotated resource to support investigators in filling this gap through translational breast cancer research. As one component of a long effort by the PRBB, the ME-BrTMA allows the opportunity to explore the tumor biology of Puerto Rican patients with breast cancer in the context of larger-scale trends in disease prognosis that we have observed.

Authors' Disclosures

M. Rosa reports grants from the NIH/NCI during the conduct of the study. J.O. Johnson reports grants from the NIH during the conduct of the study. H.I. Saavedra reports grants from R01 CA266046 and U54 CA163071 during the conduct of the study. S.A. Eschrich reports grants from the NCI/NIH during the conduct of the study and other support from Cvergenx, Inc. outside the submitted work. W.D. Cress reports grants from the NCI and nonfinancial support from the National Program of Cancer Registries during the conduct of the study. No disclosures were reported by the other authors.

Authors' Contributions

A.E. Lantz: Conceptualization, formal analysis, investigation, visualization, writing—original draft, writing—review and editing. **E.R. Gordián:** Conceptualization, resources, data curation, supervision, investigation, writing—original draft, writing—review and editing. **M. Rosa:** Resources, data curation, investigation, writing—review and editing. **M. Rodríguez Ruíz:** Data curation, investigation, writing—original draft, writing—review and editing. **J.O. Johnson:** Data curation, formal analysis, investigation, writing—review and editing. **R. Gebert:** Data curation, formal analysis, validation, project administration, writing—review and editing. **A. Bahr:** Data curation, investigation, project administration, writing—review and editing. **D.T. Chen:** Formal analysis, investigation, project administration, writing—review and editing. **J. Dutil:** Data curation, funding acquisition, investigation, project administration, writing—review and editing. **J. Li:** Formal analysis,

investigation, project administration, writing—review and editing. **J.A. Oliveras Torres:** Resources, data curation, formal analysis, supervision, investigation, project administration, writing—review and editing. **H.I. Saavedra:** Funding acquisition, investigation, project administration, writing—review and editing. **S.A. Eschrich:** Formal analysis, investigation, project administration, writing—review and editing. **I. Flores:** Conceptualization, resources, data curation, supervision, funding acquisition, investigation, project administration, writing—review and editing. **W.D. Cress:** Conceptualization, resources, data curation, formal analysis, supervision, funding acquisition, investigation, writing—original draft, project administration, writing—review and editing.

Acknowledgments

We are grateful to the anonymous patients and caregivers who contributed to this work. We acknowledge personnel at Ponce Health Sciences University (Sonia Abac, Maria Rojas, Moraima Oquendo, and Zindie Rodriguez) and the MCC (Doug Marchion, Michelle Fournier, Zena Sayegh, and Noel Clark) who made significant contributions to data collection and materials necessary to construct the TMA. Grants from the NCI (P30-CA076292, P20-CA202920, R01CA266046, U54-CA163071, and U54-CA163068) and the National Program of Cancer Registries (No. NU58DP007164) supported this work. Editorial assistance was provided by the MCC's Office of Scientific Publishing by Gerard Hebert; no compensation was given beyond his regular salary. Finally, we acknowledge Teresita Muñoz-Antonia (retired) as the founding codirector of the PRBB and the inspiration for this work.

Note

Supplementary data for this article are available at Cancer Research Communications Online (<https://aacrjournals.org/cancerrescommun/>).

Received January 09, 2025; revised March 20, 2025; accepted April 28, 2025; posted first May 01, 2025.

References

1. Spratt DE, Chan T, Waldron L, Speers C, Feng FY, Ogunwobi OO, et al. Racial/ethnic disparities in genomic sequencing. *JAMA Oncol* 2016;2:1070–4.
2. Johnson JA, Moore BJ, Syrioti G, Eden CM, Wright D, Newman LA. Landmark series: the Cancer Genome Atlas and the study of breast cancer disparities. *Ann Surg Oncol* 2023;30:6427–40.
3. Joama S, Hahn MJ, Hindorff LA, Bonham VL. Defining and achieving health equity in genomic medicine. *Ethn Dis* 2019;29(Suppl 1):173–8.
4. Landry LG, Ali N, Williams DR, Rehm HL, Bonham VL. Lack of diversity in genomic databases is a barrier to translating precision medicine research into practice. *Health Aff (Millwood)* 2018;37:780–5.
5. Lee SS, Fullerton SM, McMahon CE, Bentz M, Saperstein A, Jeske M, et al. Targeting representation: interpreting calls for diversity in precision medicine research. *Yale J Biol Med* 2022;95:317–26.
6. Sirugo G, Williams SM, Tishkoff SA. The missing diversity in human genetic studies. *Cell* 2019;177:1080.
7. Popejoy AB, Ritter DI, Crooks K, Currey E, Fullerton SM, Hindorff LA, et al; Clinical Genome Resource (ClinGen) Ancestry and Diversity Working Group (ADWG). The clinical imperative for inclusivity: race, ethnicity, and ancestry (REA) in genomics. *Hum Mutat* 2018;39:1713–20.
8. O'Donnell PH, Dolan ME. Cancer pharmacoethnicity: ethnic differences in susceptibility to the effects of chemotherapy. *Clin Cancer Res* 2009;15:4806–14.
9. Pathak S, Zajac KK, Annaji M, Govindarajulu M, Nadar RM, Bowen D, et al. Clinical outcomes of chemotherapy in cancer patients with different ethnicities. *Cancer Rep (Hoboken)* 2023;6(Suppl 1):e1830.
10. Turner BE, Steinberg JR, Weeks BT, Rodriguez F, Cullen MR. Race/ethnicity reporting and representation in US clinical trials: a cohort study. *Lancet Reg Health Am* 2022;11:100252.
11. Halmai NB, Carvajal-Carmona LG. Diversifying preclinical research tools: expanding patient-derived models to address cancer health disparities. *Trends Cancer* 2022;8:291–4.
12. Lott PC, Chiu K, Quino JE, Vang AP, Lloyd MW, Srivastava A, et al; PDXNet Consortium. Development and application of genetic ancestry reconstruction methods to study diversity of patient-derived models in the NCI PDXNet consortium. *Cancer Res Commun* 2024;4:2147–52.
13. Seijo E, Lima D, Iriabho E, Almeida J, Monico J, Echeverri M, et al. Construction and validation of a multi-institutional tissue microarray of invasive ductal carcinoma from racially and ethnically diverse populations. *Cancer Control* 2016;23:383–9.
14. Flores I, Muñoz-Antonia T, Matta J, García M, Fenstermacher D, Gutierrez S, et al. The establishment of the first cancer tissue Biobank at a Hispanic-Serving Institution: a National Cancer Institute-Funded Initiative between Moffitt cancer center in Florida and the Ponce School of Medicine and Health Sciences in Puerto Rico. *Biopreserv Biobank* 2011;9:363–71.
15. Guerrero S, López-Cortés A, Indacochea A, García-Cárdenas JM, Zambrano AK, Cabrera-Andrade A, et al. Analysis of racial/ethnic representation in select basic and applied cancer research studies. *Sci Rep* 2018;8:13978.
16. Lorona NC, Malone KE, Li CI. Racial/ethnic disparities in risk of breast cancer mortality by molecular subtype and stage at diagnosis. *Breast Cancer Res Treat* 2021;190:549–58.
17. Stringer-Reasor EM, Elkhanany A, Khoury K, Simon MA, Newman LA. Disparities in breast cancer associated with African American identity. *Am Soc Clin Oncol Educ Book* 2021;41:e29–46.
18. Stabellini N, Cullen J, Cao L, Shanahan J, Hamerschlak N, Waite K, et al. Racial disparities in breast cancer treatment patterns and treatment related adverse events. *Sci Rep* 2023;13:1233.
19. Gimbrone NT, Sarcar B, Gordian ER, Rivera JI, Lopez C, Yoder SJ, et al. Somatic mutations and ancestry markers in Hispanic lung cancer patients. *J Thorac Oncol* 2017;12:1851–6.
20. Carrot-Zhang J, Soca-Chafre G, Patterson N, Thorner AR, Nag A, Watson J, et al. Genetic ancestry contributes to somatic mutations in lung cancers from admixed Latin American populations. *Cancer Discov* 2021;11:591–8.
21. Carvajal-Carmona LG. Molecular subtypes and driver mutations in Latinos with gastric cancer: implications for etiological and translational research. In: Ramirez AG, Trapido EJ, editors. *Advancing the Science of Cancer in Latinos*. Cham (CH): Springer Nature; 2020. pp. 89–94.
22. Marker KM, Zavala VA, Vidaurre T, Lott PC, Vásquez JN, Casavilca-Zambrano S, et al; COLUMBUS Consortium. Human epidermal growth factor receptor 2-positive breast cancer is associated with indigenous American ancestry in Latin American women. *Cancer Res* 2020;80:1893–901.
23. Caliebe A, Tekola-Ayele F, Darst BF, Wang X, Song YE, Gui J, et al; IGES ELSI Committee. Including diverse and admixed populations in genetic epidemiology research. *Genet Epidemiol* 2022;46:347–71.
24. De Oliveira TC, Secolin R, Lopes-Cendes I. A review of ancestrality and admixture in Latin America and the Caribbean focusing on native American and African descendant populations. *Front Genet* 2023;14:1091269.
25. Mendez Rodríguez S, Rodríguez-Hernández AM, Torres-Torres G, Centeno-Girona H, Cruz-Correa M. Tumor molecular profiling in Hispanics: moving towards precision Oncology and health equity. *J Racial Ethn Health Disparities* 2023;10:1423–31.
26. Ding YC, Song H, Adamson AW, Schmolze D, Hu D, Huntsman S, et al. Profiling the somatic mutational landscape of breast tumors from Hispanic/Latina women reveals conserved and unique characteristics. *Cancer Res* 2023;83:2600–13.
27. Fain KM, Nelson JT, Tse T, Williams RJ. Race and ethnicity reporting for clinical trials in ClinicalTrials.gov and publications. *Contemp Clin Trials* 2021;101:106237.
28. Liu J, Lichtenberg T, Hoadley KA, Poisson LM, Lazar AJ, Cherniack AD, et al; Cancer Genome Atlas Research Network. An integrated TCGA pan-cancer clinical data resource to drive high-quality survival outcome analytics. *Cell* 2018;173:400–16.e11.
29. Wang X, Steensma JT, Bailey MH, Feng Q, Padda H, Johnson KJ. Characteristics of the Cancer Genome Atlas cases relative to U.S. general population cancer cases. *Br J Cancer* 2018;119:885–92.
30. Iqbal J, Ginsburg O, Rochon PA, Sun P, Narod SA. Differences in breast cancer stage at diagnosis and cancer-specific survival by race and ethnicity in the United States. *JAMA* 2015;313:165–73.
31. Ooi SL, Martinez ME, Li CI. Disparities in breast cancer characteristics and outcomes by race/ethnicity. *Breast Cancer Res Treat* 2011;127:729–38.
32. Akinyemiju T, Moore JX, Ojesina AI, Waterbor JW, Altekruze SF. Racial disparities in individual breast cancer outcomes by hormone-receptor subtype, area-level socio-economic status and healthcare resources. *Breast Cancer Res Treat* 2016;157:575–86.
33. Lund MJ, Trivers KF, Porter PL, Coates RJ, Leyland-Jones B, Brawley OW, et al. Race and triple negative threats to breast cancer survival: a population-based study in Atlanta, GA. *Breast Cancer Res Treat* 2009;113:357–70.
34. Lund MJ, Butler EN, Bumpers HL, Okoli J, Rizzo M, Hatchett N, et al. High prevalence of triple-negative tumors in an urban cancer center. *Cancer* 2008;113:608–15.
35. Martínez ME, Gomez SL, Tao L, Cress R, Rodriguez D, Unkart J, et al. Contribution of clinical and socioeconomic factors to differences in breast cancer subtype and mortality between Hispanic and non-Hispanic white women. *Breast Cancer Res Treat* 2017;166:185–93.
36. Champion CD, Thomas SM, Plichta JK, Parrilla Castellar E, Rosenberger LH, Greenup RA, 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:e827–38.
37. Miller KD, Goding Sauer A, Ortiz AP, Fedewa SA, Pinheiro PS, Tortolero-Luna G, et al. Cancer statistics for Hispanics/Latinos, 2018. *CA Cancer J Clin* 2018;68:425–45.
38. Huo D, Ikpat F, Khramtsov A, Dangou JM, Nanda R, Dignam J, et al. Population differences in breast cancer: survey in indigenous African women reveals over-representation of triple-negative breast cancer. *J Clin Oncol* 2009;27:4515–21.

39. Giaquinto AN, Miller KD, Tossas KY, Winn RA, Jemal A, Siegel RL. Cancer statistics for African American/Black people 2022. *CA Cancer J Clin* 2022;72:202–29.
40. Ortiz AP, Frias O, Pérez J, Cabanillas F, Martínez L, Sánchez C, et al. Breast cancer molecular subtypes and survival in a hospital-based sample in Puerto Rico. *Cancer Med* 2013;2:343–50.
41. Amirikia KC, Mills P, Bush J, Newman LA. Higher population-based incidence rates of triple-negative breast cancer among young African-American women: implications for breast cancer screening recommendations. *Cancer* 2011;117:2747–53.
42. Stark A, Kleer CG, Martin I, Awuah B, Nsiah-Asare A, Takyi V, et al. African ancestry and higher prevalence of triple-negative breast cancer: findings from an international study. *Cancer* 2010;116:4926–32.
43. Hines LM, Risendal B, Byers T, Mengshol S, Lowery J, Singh M. Ethnic disparities in breast tumor phenotypic subtypes in Hispanic and non-Hispanic white women. *J Womens Health (Larchmt)* 2011;20:1543–50.
44. Howlader N, Altekruse SF, Li CI, Chen VW, Clarke CA, Ries LA, et al. US incidence of breast cancer subtypes defined by joint hormone receptor and HER2 status. *J Natl Cancer Inst* 2014;106:dju055.
45. Banegas MP, Li CI. Breast cancer characteristics and outcomes among Hispanic Black and Hispanic White women. *Breast Cancer Res Treat* 2012;134:1297–304.
46. Sineshaw HM, Gaudet M, Ward EM, Flanders WD, Desantis C, Lin CC, et al. Association of race/ethnicity, socioeconomic status, and breast cancer subtypes in the National Cancer Data Base (2010–2011). *Breast Cancer Res Treat* 2014;145:753–63.
47. Chen L, Li CI. Racial disparities in breast cancer diagnosis and treatment by hormone receptor and HER2 status. *Cancer Epidemiol Biomarkers Prev* 2015;24:1666–72.
48. D'Arcy M, Fleming J, Robinson WR, Kirk EL, Perou CM, Troester MA. Race-associated biological differences among Luminal A breast tumors. *Breast Cancer Res Treat* 2015;152:437–48.
49. Goel N, Yadegarynia S, Lubarsky M, Choi S, Kelly K, Balise R, et al. Racial and ethnic disparities in breast cancer survival: emergence of a clinically distinct Hispanic Black population. *Ann Surg* 2021;274:e269–75.
50. Torres-Cintrón CR, Suárez-Ramos T, Román-Ruiz Y, Ortiz-Ortiz KJ, De Jesús-Monge V, Gierbolini-Bermúdez A, et al. Cancer in Puerto Rico 2016–2020. San Juan (PR): Puerto Rico Central Cancer Registry; 2023.
51. Morales L, Alvarez-Garriga C, Matta J, Ortiz C, Vergne Y, Vargas W, et al. Factors associated with breast cancer in Puerto Rican women. *J Epidemiol Glob Health* 2013;3:205–15.
52. Serrano-Gómez SJ, Fejerman L, Zabaleta J. Breast cancer in Latinas: a focus on intrinsic subtypes distribution. *Cancer Epidemiol Biomarkers Prev* 2018;27:3–10.
53. Lantz AE, Gebert R, Li J, Oliveras JA, Gordián ER, Perez-Morales J, et al. Worse clinical and survival outcomes in breast cancer patients living in Puerto Rico compared to Hispanics, non-Hispanic Blacks, and non-Hispanic Whites from Florida. *J Racial Ethn Health Disparities* 2024 Nov 14 [Epub ahead of print].
54. Wolff AC, Hammond MEH, Allison KH, Harvey BE, Mangu PB, Bartlett JMS, et al. Human epidermal growth factor receptor 2 testing in breast cancer: American Society of Clinical Oncology/College of American pathologists clinical practice guideline focused update. *J Clin Oncol* 2018;36:2105–22.
55. Allison KH, Hammond MEH, Dowsett M, McKernin SE, Carey LA, Fitzgibbons PL, et al. Estrogen and progesterone receptor testing in breast cancer: American Society of Clinical Oncology/College of American pathologists guideline update. *Arch Pathol Lab Med* 2020;144:545–63.
56. Gendoo DM, Ratanasirigulchai N, Schröder MS, Paré L, Parker JS, Prat A, et al. Genefu: an R/Bioconductor package for computation of gene expression-based signatures in breast cancer. *Bioinformatics* 2016;32:1097–9.
57. Bornstein-Quevedo L, de Anda-González J, Lara-Torres CO, Flores-Gutiérrez JP, Dorantes-Heredia R, Bautista-Piña V, et al. Navigating HER2-low testing in invasive breast cancer: update recommendations for pathologists. *J Pers Med* 2024;14:467.
58. Feeley LP, Mulligan AM, Pinnaduwa D, Bull SB, Andrulis IL. Distinguishing luminal breast cancer subtypes by Ki67, progesterone receptor or TP53 status provides prognostic information. *Mod Pathol* 2014;27:554–61.
59. Nielsen TO, Leung SCY, Rimm DL, Dodson A, Acs B, Badve S, et al. Assessment of Ki67 in breast cancer: updated recommendations from the international Ki67 in breast cancer working group. *J Natl Cancer Inst* 2021;113:808–19.
60. Badowska-Kozakiewicz AM, Budzik MP. Immunohistochemical characteristics of basal-like breast cancer. *Contemp Oncol (Pozn)* 2016;20:436–43.
61. Rivera-Rivera Y, Vargas G, Jaiswal N, Núñez-Marrero A, Li J, Chen DT, et al. Ethnic and racial-specific differences in levels of centrosome-associated mitotic kinases, proliferative and epithelial-to-mesenchymal markers in breast cancers. *Cell Div* 2022;17:6.
62. Swain SM, Shastry M, Hamilton E. Targeting HER2-positive breast cancer: advances and future directions. *Nat Rev Drug Discov* 2023;22:101–26.
63. Ma Q, Liu YB, She T, Liu XL. The role of Ki-67 in HR+/HER2- breast cancer: a real-world study of 956 patients. *Breast Cancer* 2024;16:117–26.
64. Davey MG, Hynes SO, Kerin MJ, Miller N, Lowery AJ. Ki-67 as a prognostic biomarker in invasive breast cancer. *Cancers (Basel)* 2021;13:4455.
65. Guth AA, Chun Kim J, Schwartz S, Montes J, Snyder RA, Axelrod D, et al. The relationship of race, Oncotype DX, and Ki67 in a population highly screened for breast cancer. *Breast J* 2017;23:177–81.