Published in final edited form as:

Breast Cancer Res Treat. 2018 November ; 172(1): 123–132. doi:10.1007/s10549-018-4879-7.

Intra-tumor molecular heterogeneity in breast cancer: definitions of measures and association with distant recurrence-free survival

Ashirbani Saha, PhD¹, Michael R. Harowicz, MD¹, Elizabeth Hope Cain, BA¹, Alison H. Hall², Eun-Sil Shelley Hwang, MD, MPH³, Jeffrey R. Marks, PhD⁴, Paul Kelly Marcom, MD⁵, and Maciej A. Mazurowski, PhD^{1,6,7}

¹Department of Radiology, Duke University School of Medicine, Durham, NC 27705, USA

²Department of Pathology, Duke University School of Medicine, Durham NC 27710, USA

³Department of Surgical Oncology, Duke University School of Medicine, Durham, NC 27710, USA

⁴Department of Surgery, Duke University School of Medicine, Durham, NC 27710, USA

⁵Department of Medicine, Duke University School of Medicine, Durham, NC 27710, USA

⁶Department of Electrical and Computer Engineering, Duke University, Durham, NC 27705, USA

⁷Duke University Medical Physics Program, Durham, NC 27705, USA

Abstract

Purpose: To define quantitative measures of intra-tumor heterogeneity in breast cancer based on histopathology data gathered from multiple samples on individual patients and determine their association with distant recurrence-free survival (DRFS).

Methods: We collected data from 971 invasive breast cancers, from 1st January 2000 to 23rd March 2014, that underwent repeat tumor sampling at our institution. We defined and calculated 31 measures of intra-tumor heterogeneity including ER, PR, and HER2 immunohistochemistry (IHC), proliferation, EGFR IHC, grade and histology. For each heterogeneity measure, Cox proportional hazards models were used to determine whether patients with heterogeneous disease had different distant recurrence-free survival (DRFS) than those with homogeneous disease.

Results: Presence of heterogeneity in ER percentage staining was prognostic of reduced DRFS with a hazard ratio of 4.26 (95% CI: 2.22– 8.18, $p < 0.00002$). It remained significant after controlling for the ER status itself ($p < 0.00062$) and for patients that had chemotherapy ($p < 0.00032$). Most of the heterogeneity measures did not show any association with DRFS despite the considerable sample size.

Corresponding Author: Ashirbani Saha, PhD, Department of Radiology, Duke University Medical Center, 2424 Erwin Road, Suite 302, Durham, NC 27705, USA, Phone: +1 919 684 1458, ashirbani.saha@duke.edu.

Conflicts of Interest

The authors have no conflicts of interest to disclose

Conclusions: Intra-tumor heterogeneity of ER receptor status may be a predictor of patient DRFS. Histopathologic data from multiple tissue samples may offer a view of tumor heterogeneity and assess recurrence risk.

Keywords

breast cancer; immunohistochemistry; distant recurrence-free survival; receptor status; intra-tumor heterogeneity; survival analysis

Introduction:

Tumor heterogeneity between patients (inter-tumor heterogeneity) in breast cancer has been described both in terms of the clinical behavior and in terms of tumor pathology/genomics [1–5]. Based on histopathologic characteristics and gene expression profiles, breast tumors can be classified into multiple categories. Prior studies have demonstrated that tumors with distinct pathologic characteristics show unique clinical behavior and should be treated differently. Specifically, receptor status (ER, PR, HER2) of breast tumor cells is an important and established measure used in therapeutic decision-making [6, 7] as the receptors are both targets for treatment as well as prognostic indicators [8–13].

More recent studies show that in addition to the inter-tumor heterogeneity, breast cancers can demonstrate important genetic and phenotypic differences within a tumor itself, a phenomenon referred to as intra-tumor heterogeneity [14–16]. This heterogeneity can manifest as morphologic and biochemical variation. Tumor heterogeneity likely arises from a combination of intrinsic cellular differences (including genetic and epigenetic changes), as well as variation in the microenvironment. This variation in tumor phenotype provides the basis for tumor evolution, with implications for critical biologic processes, including disease progression and resistance to therapy.

Intra-tumor heterogeneity in breast cancer remains relatively unexplored (compared to inter-tumor heterogeneity), partially due to the challenges in quantitatively characterizing and measuring heterogeneity. Previous studies have showed that histopathologic samples of different sites of the tumor may produce varying characteristics [17–20]. In other studies [21–23], investigators have proposed measures to quantify HER2 heterogeneity to assess regional and genetic HER2 heterogeneity, or cell level and tumor level HER2 heterogeneity. Heterogeneity in intra-tumoral ER/PR/EGFR immunohistochemistry (IHC) expression has also been analyzed [24–26]. Specifically, these three studies focused on different aspects of heterogeneity as follows: (a) cell populations having different distributions of ER, PR and cytokeratin 5 in a single tumor and a sub-population of cells (ER,-/PR,-/CK5+) were identified for additional treatment [24]; (b) the distribution of ER status in different regions of the breast cancer [25]; and (c) prognosis of distant metastasis using tumor heterogeneity of ER patterns in contralateral breast cancer [26].

Because of intra-tumoral heterogeneity, a single sample of a tumor may not be sufficient to provide a complete characterization of genetic, epigenetic, and/or phenotypic changes of the whole tumor [1]. For example, in a newly diagnosed breast cancer, biomarker analysis (immunohistochemistry) is based on a single small sample of what is frequently a much

larger tumor. As breast biomarker studies play a key role in determining the treatment course, heterogeneity in these factors can complicate therapeutic decision making [27]. A recent study [28] recommends implementation of intra-tumor heterogeneity measures of breast biomarkers to guide treatment-related decision making.

Additionally, the degree of genetic variation may be even greater than the degree of heterogeneity in protein expression. Despite having similar receptor expression, multifocal breast cancers have been shown to exhibit genomic heterogeneity [29]. Genomic intra-tumor heterogeneity has strong clinical implications [30] and its presence in breast cancer is suspected to be associated with development of treatment resistance. In addition to these, imaging markers of intra-tumor heterogeneity are beginning to emerge [31].

These studies demonstrate that a variety of approaches have been taken to measure or detect the presence of intra-tumor heterogeneity and analyze its effects. It is important to further explore ways to measure intra-tumor heterogeneity and subsequently analyze which measurements are associated with clinical outcomes. In this study, we address both of these issues in the context of standard clinical pathology reports. The first objective of this study was to develop quantitative measures of tumor heterogeneity using components of the IHC data from multiple breast tumor samples. Though this approach has been studied earlier [26] our study is broader as it utilizes numerous different IHC components and has a expansive patient population as the aforementioned study was restricted to patients with contralateral breast cancer. Moreover, our second objective was assessing the association between the proposed heterogeneity measures and patient outcomes, here defined as DRFS. A recent study [28] reports that histologic data from different samples provides an increased view of breast tumor biology. Despite this, its subsequent usage in the clinical setting for diagnosis, prognosis, and therapy has not gained much momentum. We performed a comprehensive analysis using heterogeneity based on IHC data and took into account the effect of three different therapies (chemotherapy, endocrine therapy, and anti-HER2 therapy) in our analysis. Throughout the study, we have used different abbreviations and they are listed in Table 1.

Materials and Methods:

Patient Population

In this retrospective, Institutional Review Board-approved study, we identified 971 patients with invasive breast cancer and pathology data sufficient to assess at least one measure of heterogeneity. The detailed selection criteria are shown in Fig. 1, which shows that patients were excluded for unavailability of pertinent data and for not satisfying some clinical conditions. Regarding surgery based exclusion, we excluded patients for the initial cohort (1166 patients) with prior non-definitive breast surgery before MRI (having mastoplasty or lumpectomy, but excluding implants) and definitive breast surgery before MRI. We followed REporting recommendations for tumor MARKer (REMARK) [32] prognostic study criteria for this study.

Data from pathology samples

For each patient, we retrieved data based on all available pathology and biomarker reports. We found that the patients in our cohort had up to 6 pathology samples available. The pathology samples were collected either from biopsy (57.11%), lumpectomy (24.03%), or mastectomy (18.86%). The median time-difference between two samples collected was 45 days. In our cohort, the distribution of patients with each number of samples was the following: 1 sample: 60 patients, 2 samples: 629 patients, 3 samples: 198 patients, 4 samples: 69 patients, 5 samples: 13 patients and 6 samples: 2 patients. The data collected included the following information: (a) date, (b) sample type (biopsy, lumpectomy or mastectomy), (c) sample side (left or right), (d) ER and PR IHC (percentage staining, intensity score, Allred score), (e) HER2 IHC (percentage staining, score, Automated cellular imaging system -ACIS [33], FISH ratio, FISH SD, percentage of cells with HER2/ chromosome 17 centromere ratio greater than 2.2, average copy of chromosome 17, average copy of HER2/neu gene), (f) Ki-67 index, (g) epidermal growth factor receptor (EGFR) IHC (percentage staining and score), (h) histological grade data (Nottingham grade score, tubule formation score, nuclear pleomorphism score, mitotic rate score), and (i) invasive carcinoma histology subtypes (ductal, lobular, micropapillary, poorly differentiated, metaplastic, medullary, tubular, mucinous, and colloid).

Heterogeneity measures

From the available pathology data, we defined and calculated 31 measures of tumor heterogeneity for each patient, as feasible. These measures were based on ER IHC (4 measures, 2 from ER percentage staining and 2 from ER Allred scores), PR IHC (4 measures, 2 from PR percentage staining and 2 from PR Allred scores), HER2 IHC (11 measures, 2 from HER2 percentage staining, 1 from HER2 scores, 1 from ACIS, 1 from FISH ratio, 2 from FISH SD, 2 from percentage of cells with HER2 FISH based copy number exceeding a ratio of 2.2, 1 from average copy of chromosome 17, 1 from average copy of HER2 neu gene), Ki-67 index (2 measures), EGFR IHC (4 measures, 2 from EGFR percentage staining and 2 from EGFR scores), tumor grade (5 measures, 1 from Nottingham grade, 1 from tubule formation score, nuclear pleomorphism score and mitotic rate score each, 1 from the combination of tubule formation, nuclear pleomorphism and mitotic rate scores) and type (1 measure). Out of these measures, 29 are based on multiple samples and require the presence of relevant data from at least 2 samples. Two out of the 31 measures - FISH SD and percentage of cells with HER2 FISH ratio greater than 2.2 are single-sample measures and are assessed based on the first available sample. The quantitative measures of heterogeneity were comprised of continuous, ternary, and binary valued features. The names, full definitions and details of all 31 measures are mentioned in the Supplementary Material (eTable 1 in Part-A).

Outcomes and therapy data

The outcome of interest in this study was distant recurrence. The time to last distant recurrence-free follow-up was collected from the patient's medical records until 3rd May 2016 as per study protocol. For two patients among the cohort of 971, the date of distant

recurrence and days to last follow-up before recurrence were not available. These two patients were excluded from the statistical analysis.

We also collected data from medical records on whether the patient received each of the following therapies: chemotherapy, endocrine therapy and anti-HER2/neu therapy. These were recorded and analyzed as binary variables.

Statistical Analysis

To determine whether each measure of heterogeneity was associated with DRFS, we used a univariate Cox proportional hazards regression model for each variable and recorded the significance of each of the associations. We used $p < 0.0016$ as the significance level based on the number of comparisons tested for our primary hypothesis ($0.05/31$ features). To compute the regression model, we turned each of the measures of heterogeneity into a binary variable to indicate the presence or absence of intra-tumor heterogeneity in the patient for each given feature. We used the hazard ratio (HR) to determine whether each of the measures of heterogeneity was associated with DRFS. In order to do so, we binarized all of the measures to classify the tumors as homogeneous or heterogeneous according to each of the measures. For continuous measures of heterogeneity, the median value of the measure was used as the cut-off for intra-tumor heterogeneity. All patients having values greater than the median were deemed the heterogeneous group, and otherwise they were considered homogeneous. For the ternary feature, values greater than 1 represented the heterogeneous group. The heterogeneity measures that were originally binary by definition did not change in the hazard ratio analysis. We also analyzed pairwise correlation for all binarized features. The available sample size for each distinct measure was variable, as each requires information from different types of pathology records.

In an additional exploratory analysis, we conducted the same analysis for the following subgroups of patients: (1) those that received chemotherapy, (2) those that received endocrine therapy, and (3) those that received anti HER2/neu therapy. This allowed us to control for the therapy regimen and tumor phenotype. Finally, for each variable, we determined whether heterogeneity in a given marker is prognostic of DRFS independently of the status of the marker itself (details can be found in Part B of the supplementary material.).

Results:

The overall demographic and clinicopathologic characteristics of our patient population are shown in Table 1. This table also presents the therapy-wise (chemotherapy, endocrine therapy, and anti-HER2/neu therapy-wise) demographic and clinicopathologic characteristics.

The distribution of the clinicopathologic characteristics for each heterogeneity measure and specifically within the heterogeneous and homogeneous sub-groups of each measure are shown in Supplementary material (eFigs. 1–31 in Part-C) provided with this paper. Please note that size of the usable cohort for each measure varies significantly and also changes when a specific therapy is considered. Please also note that the size of cohort obtained for measures calculated from continuous marker values (e.g. ER percentage staining) and

corresponding binary marker values may be different due to a varying level of detail in a pathology report. At times, the value of one binary variable was established but not the continuous variable (e.g., “positive”). Please refer to the supplementary material (eTable 1, column ‘Definition’) for additional details.

The correlations between measures after binarizing them are presented in Fig. 2.

Figure 2 shows the correlation between a pair of heterogeneity measures (after binarization, when required). As expected, we found that measures computed using the same type of information (e.g. same marker status related, grade related) are more correlated with each other than with other variable groups. The binary and continuous features from ER Allred scores are correlated with the rest of the features of ER IHC. A similar observation is made from PR IHC. For HER2 IHC, we observe two distinct groups that are more correlated among themselves. In the first group, the features from HER2 percentage staining, HER2 score, and ACIS are present. In the second group, features from FISH ratio, FISH SD, and MDBS cont. Avg copy of neu gene are present. The features from Ki67 are correlated with each other. Also, features from EGFR IHC have a high correlation within the group. The MDBS cont. combined score from grade feature, which quantifies the maximum cumulative difference in tubule formation score, nuclear pleomorphism score, and mitotic rate score, is highly correlated with rest of the features from grade but is less correlated with all other features. MDBS cont. ER Allred score and FISH ratio from first sample greater than 2.2 have high correlations with features from other types as well. The correlations of the features within same and different IHC groups are shown in Table 2, which also summarizes the findings from Fig. 2. This also shows that heterogeneity measures from different IHC related groups complement each other more compared to the measures from the same group. As different clinical markers of IHC play significant roles in the management of the disease [34, 35], the complementary information in the heterogeneity of the marker may contain significant information for the management of the disease.

Figure 3 shows the prognostic value (HR) of predicting DRFS for each of the heterogeneity measures. This is demonstrated for the entire sample available for each feature as well as for patients that received specific types of therapy. Out of 31 features, 7 features were found for the available population such that the HR values did not overlap with 1. For chemotherapy and endocrine therapy, 5 such features were found. As seen in the figure, the measure MDBS binary ER percent staining, which captures the heterogeneity in terms of discordance between binarized ER percentage staining values, showed a high HR of 4.26 (95% CI: 2.22–8.18) for distant recurrence of cancer. This measure is a strongly significant predictor of recurrence with p-value less than 0.00002. This measure of heterogeneity remained significant after controlling for chemotherapy ($p < 0.00032$) with an HR of 3.38 (95% CI: 1.74–6.54). Controlling for endocrine therapy resulted in an HR of 4.12 (95% CI: 1.25–13.65) and p-value of 0.0203. There were no events/distant recurrences recorded in the heterogeneous subset when controlled for anti-HER2 therapy for this measure. The Kaplan-Meier plots for these tests are shown in Fig. 4. As shown in this figure, for available population, chemotherapy, and endocrine therapy more patients in heterogeneous groups have a distant recurrence compared to the homogeneous groups. In the anti-HER2 patients,

no recurrence was found in heterogeneous groups resulting in 95% confidence interval for HR between 0 and infinity for the measure MDBS binary ER percent staining.

Two more measures showed significant prognostic capability when controlled for endocrine therapy. The first of these is MDBS binary ER Allred score, which captures the presence of discordance in the binarized ER Allred scores between samples. It showed an HR of 5.96 (95% CI: 2.08–17.08) with p-value of 0.000891. The second statistically significant measure is MDBS binary PR percent staining, which computes tumor heterogeneity from binarized PR percentage staining values. It was significant with an HR of 3.57 (95% CI: 1.63–7.86) and a p-value of 0.00152. The Kaplan-Meier plots for each measure of heterogeneity can be found in Supplementary material (eFigs.32–62 in Part-D).

For all available population, the measure MDBS binary ER percent staining remained a significant predictor ($p < 0.00062$) of DRFS when controlled for ER status with HR of 3.43 (95% CI: 1.69–6.94). Among the measures from PR IHC, MDBS binary PR percent staining remained a significant predictor ($p < 0.0012$) of DRFS after controlling for PR status with HR of 2.86 (95% CI: 1.51 – 5.39).

Discussion:

In this study, we have defined and computed 31 different measures to quantify breast tumor heterogeneity based on commonly collected pathologic and molecular parameters. The investigation was based on the concept that greater heterogeneity may promote tumor growth and resistance to targeted therapy.

The proposed measures of heterogeneity are diverse in source as they are gathered from pathologic evaluation and common tumor markers through immunohistochemistry and FISH. They are also diverse in their range and type of value. Continuous and ternary variables quantify the degree of heterogeneity, while binary measures provide a direct indication of breast tumor heterogeneity. From the correlation plot in Fig. 2 and Table 3, we find that measures computed using the same type of information (for example: ER IHC based measures) are more correlated. Measures obtained using different types of histologic data have unique information to offer.

Using one measure of heterogeneity at a time, we divided the patient cohort into two groups: heterogeneous and homogeneous. Then we analyzed if the heterogeneous group of patients have an increased rate of distant recurrence. The measure showing discordance between binarized ER percentage staining of multiple samples demonstrated a very strong association with a higher risk of distant metastasis. This was the case for the entire available population and the population treated with chemotherapy. Two other binary measures computed using ER and PR IHC data showed promising prognostic capability. The fact that more binary measures showed promising prognostic ability than continuous measures, indicates that binarization with a proper threshold is an important factor in detecting heterogeneity associated with a higher risk of recurrence. There are variations in terms of thresholds/cut-off values for the determination of receptor status in the literature [36]. Choosing a specific threshold for a given binary measure is critical because this defines heterogeneous versus

homogeneous tumor and can affect treatment decisions. Based on the strong performance of binarized measures computed using percentage staining of ER, PR, and EGFR, we can conclude that the threshold used in our study is a good choice for determining the risk of distant recurrence in our cohort.

We found two binary measures of heterogeneity to be significantly prognostic of DRFS after controlling for ER and PR status: ER percentage staining and PR percentage staining. This shows that heterogeneity in ER and PR percentage staining can be prognostic of DRFS, independent of the marker status itself.

Three continuous measures showed promising performance in separating the available population into high-risk and low-risk groups. These measures are based on the difference between the continuous PR Allred score, the difference between the continuous EGFR scores and the maximum difference in mitotic scores of tumor grade between samples. In our analysis, thresholding of the continuous values was driven by the median value of the measure from the available cohort. While the continuous measures quantify the degree of heterogeneity for the particular IHC characteristic, thresholding in this particular manner didn't prove to be effective for other continuous measures. Alternate ways of establishing a meaningful threshold should be explored in future.

None of the measures demonstrated statistically significant prognostic value when analyzed within the anti-HER2/neu therapy cohort. As demonstrated by the clinicopathologic characteristics in the supplementary material (eFigs.1–31 in Part-A), the patient population and number of recurrences are significantly smaller for this therapy, which renders the inconclusive tests for this therapy group.

Our study had some limitations. The study was performed at one institution. The data gathered spans over 14 years, which allows for some variability in the analysis of pathology slides. The study is retrospective in nature. However, the population represents variation in age, race, tumor staging, menopausal status, and in the presence of multi-focal/multi-centric tumors. We were also able to control for different therapies received. Our study is restricted to patients having a pre-operative MRI. Thus, patients at higher risk may be slightly overrepresented in the study. The rate of distant recurrences was 8.75% for our patient cohort with a median follow-up time of 4 years. However, the higher risk population is of particular interest in this context. Our results show that the actual proportion of metastasis among higher risk patients can be further determined by analyzing tumor heterogeneity using multiple samples and therefore it contributes to a better precision. For the majority of patients, staining was done at the same institution. Hence, the data is less affected by institutional bias.

Importantly, we did not account for the interobserver variability/technique induced variability which could be a potential source of heterogeneity between samples. Therefore, the true differences in tumor genomics can be confounded by these two factors. However, a review of the prior literature [37, 38] shows that heterogeneity found in our study is notably higher than found in the studies related to observer variability. Therefore, a minor part of the variability comes from the interobserver/technique induced variability and the majority of

the differences are from inter-sample variability. For example, in our study, 8.78% of average discordance between a pair of samples for the patients was observed using positivity/negativity for ER percentage staining which is greater than the discordance (<5%) seen between the observers in the study by Reisenbichler et al. [38]. Moreover, using ER Allred score, a difference of more than 1 point in 20.62% percent of the cases was found whereas Reisenbichler et al. found that occurring in less than 10% of their cases. Overall in our study, the strong effect of the heterogeneity in binarized ER percentage staining values from different samples persisted when controlled for chemotherapy and the receptor status itself.

In conclusion, we proposed detailed definitions of intra-tumor heterogeneity given multiple standard of care pathology samples and determined which measures might be associated with poor DRFS. Future studies could validate these measures at other institutions as well as explore their applicability for predicting DRFS in other cohorts.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgements

We gratefully acknowledge funding from the National Institutes of Health, 1R01EB021360 (MM), 5R01CA185138-01 (ESH), and the North Carolina Biotechnology Center, 2016-BIG-6520, (MM).

References

1. Martelotto LG, Ng CKY, Piscuoglio S, Weigelt B, and Reis-Filho JS, Breast cancer intra-tumor heterogeneity. *Breast Cancer Research*, 2014 16(3).
2. Hunter K, The role of individual inheritance in tumor progression and metastasis. *Journal of Molecular Medicine*, 2015 93(7): p. 719–725. [PubMed: 26054921]
3. Davnall F, Yip CSP, Ljungqvist G, et al., Assessment of tumor heterogeneity: An emerging imaging tool for clinical practice? *Insights into Imaging*, 2012 3(6): p. 573–589. [PubMed: 23093486]
4. Bedard PL, Hansen AR, Ratain MJ, and Siu LL, Tumour heterogeneity in the clinic. *Nature*, 2013 501(7467): p. 355–364. [PubMed: 24048068]
5. Henderson IC and Patek AJ, The relationship between prognostic and predictive factors in the management of breast cancer. *Breast Cancer Research and Treatment*, 1998 52(1): p. 261–288. [PubMed: 10066087]
6. Esserman LJ, Berry DA, Cheang MCU, et al., Chemotherapy response and recurrence-free survival in Neoadjuvant breast cancer depends on biomarker profiles: Results from the I-SPY 1 TRIAL (CALGB 150007/150012; ACRIN 6657). *Breast Cancer Research and Treatment*, 2012 132(3): p. 1049–1062. [PubMed: 22198468]
7. Walker RA, Immunohistochemical markers as predictive tools for breast cancer. *Journal of clinical pathology*, 2008 61(6): p. 689–696. [PubMed: 18037665]
8. Savci-Heijink CD, Halfwerk H, Hooijer GJK, et al., Retrospective analysis of metastatic behaviour of breast cancer subtypes. *Breast Cancer Research and Treatment*, 2015 150(3): p. 547–557. [PubMed: 25820592]
9. Bae SY, Kim S, Lee JH, et al., Poor prognosis of single hormone receptor-positive breast cancer: similar outcome as triple-negative breast cancer. *BMC Cancer*, 2015 15(1): p. 138. [PubMed: 25880075]

10. Slamon DJ, Clark GM, Wong SG, et al., Human breast cancer: correlation of relapse and survival with amplification of the HER-2/neu oncogene. *Science*, 1987 235(4785): p. 177. [PubMed: 3798106]
11. Seol H, Lee HJ, Choi Y, et al., Intratumoral heterogeneity of HER2 gene amplification in breast cancer: its clinicopathological significance. *Mod Pathol*, 2012 25(7): p. 938–948. [PubMed: 22388760]
12. Slamon DJ, Leyland-Jones B, Shak S, et al., Use of chemotherapy plus a monoclonal antibody against HER2 for metastatic breast cancer that overexpresses HER2. *New England Journal of Medicine*, 2001 344(11): p. 783–792. [PubMed: 11248153]
13. Romond EH, Perez EA, Bryant J, et al., Trastuzumab plus adjuvant chemotherapy for operable HER2-positive breast cancer. *New England Journal of Medicine*, 2005 353(16): p. 1673–1684. [PubMed: 16236738]
14. Marusyk A, Almendro V, and Polyak K, Intra-tumour heterogeneity: a looking glass for cancer? *Nature Reviews Cancer*, 2012 12(5): p. 323–334. [PubMed: 22513401]
15. Ng CKY, Pemberton HN, and Reis-Filho JS, Breast cancer intratumor genetic heterogeneity: causes and implications. *Expert review of anticancer therapy*, 2014.
16. Swanton C, Intratumor heterogeneity: evolution through space and time. *Cancer research*, 2012 72(19): p. 4875–4882. [PubMed: 23002210]
17. Davis BW, Zava DT, Locher GW, Goldhirsch A, and Hartmann WH, Receptor heterogeneity of human breast cancer as measured by multiple intratumoral assays of estrogen and progesterone receptor. *European Journal of Cancer and Clinical Oncology*, 1984 20(3): p. 375–382. [PubMed: 6323188]
18. Pertschuk LP, Axiotis CA, Feldman JG, et al., Marked Intratumoral Heterogeneity of the Proto-Oncogene Her-2/neu Determined by Three Different Detection Systems. *The breast journal*, 1999 5(6): p. 369–374. [PubMed: 11348316]
19. Nassar A, Radhakrishnan A, Cabrero IA, Cotsonis GA, and Cohen C, Intratumoral heterogeneity of immunohistochemical marker expression in breast carcinoma: a tissue microarray-based study. *Applied Immunohistochemistry & Molecular Morphology*, 2010 18(5): p. 433–441. [PubMed: 20485156]
20. Wilking U, Karlsson E, Skoog L, et al., HER2 status in a population-derived breast cancer cohort: discordances during tumor progression. *Breast cancer research and treatment*, 2011 125(2): p. 553–561. [PubMed: 20628810]
21. Buckley NE, Forde C, McArt DG, et al., Quantification of HER2 heterogeneity in breast cancer-implications for identification of sub-dominant clones for personalised treatment. *Scientific Reports*, 2016 6: p. 23383. [PubMed: 26996207]
22. Potts SJ, Krueger JS, Landis ND, et al., Evaluating tumor heterogeneity in immunohistochemistry-stained breast cancer tissue. *Laboratory Investigation*, 2012 92(9): p. 1342–1357. [PubMed: 22801299]
23. Seol H, Lee HJ, Choi Y, et al., Intratumoral heterogeneity of HER2 gene amplification in breast cancer: its clinicopathological significance. *Modern Pathology*, 2012 25(7): p. 938–948. [PubMed: 22388760]
24. Knox AJ, Scaling AL, Pinto MP, et al., Modeling luminal breast cancer heterogeneity: combination therapy to suppress a hormone receptor-negative, cytokeratin 5-positive subpopulation in luminal disease. *Breast Cancer Research*, 2014 16(4): p. 1.
25. van Netten JP, Armstrong JB, Carlyle SS, et al., Estrogen receptor distribution in the peripheral, intermediate and central regions of breast cancers. *European Journal of Cancer and Clinical Oncology*, 1988 24(12): p. 1885–1889. [PubMed: 3220085]
26. Sandberg MEC, Hartman M, Klevebring D, et al., Prognostic implications of estrogen receptor pattern of both tumors in contralateral breast cancer. *Breast cancer research and treatment*, 2012 134(2): p. 793–800. [PubMed: 22622811]
27. Koren S and Bentires-Alj M, Breast Tumor Heterogeneity: Source of Fitness, Hurdle for Therapy. *Molecular Cell*, 2015 60(4): p. 537–546. [PubMed: 26590713]

28. Beca F and Polyak K, Intratumor Heterogeneity in Breast Cancer, in Novel Biomarkers in the Continuum of Breast Cancer, Stearns V, Editor. 2016, Springer International Publishing: Cham p. 169–189.
29. Desmedt C, Fumagalli D, Pietri E, et al., Uncovering the genomic heterogeneity of multifocal breast cancer. *The Journal of pathology*, 2015 236(4): p. 457–466. [PubMed: 25850943]
30. McGranahan N and Swanton C, Biological and Therapeutic Impact of Intratumor Heterogeneity in Cancer Evolution. *Cancer Cell*. 27(1): p. 15–26. [PubMed: 25584892]
31. Chaudhury B, Zhou M, Goldgof DB, et al., Heterogeneity in intratumoral regions with rapid gadolinium washout correlates with estrogen receptor status and nodal metastasis. *Journal of Magnetic Resonance Imaging*, 2015 42(5): p. 1421–1430. [PubMed: 25884277]
32. McShane LM, Altman DG, Sauerbrei W, et al., REporting recommendations for tumour MARKer prognostic studies (REMARK). *Br J Cancer*, 2005 93(4): p. 387–391. [PubMed: 16106245]
33. Wang S, Saboorian MH, Frenkel EP, et al., Assessment of HER-2/neu status in breast cancer. *American journal of clinical pathology*, 2001 116(4): p. 495–503. [PubMed: 11601134]
34. Spitale A, Mazzola P, Soldini D, Mazzucchelli L, and Bordoni A, Breast cancer classification according to immunohistochemical markers: clinicopathologic features and short-term survival analysis in a population-based study from the South of Switzerland. *Annals of Oncology*, 2009 20(4): p. 628–635. [PubMed: 19074747]
35. Subik K, Lee J-F, Baxter L, et al., The Expression Patterns of ER, PR, HER2, CK5/6, EGFR, Ki-67 and AR by Immunohistochemical Analysis in Breast Cancer Cell Lines. *Breast Cancer: Basic and Clinical Research*, 2010 4: p. 117822341000400004.
36. Chen X, Yuan Y, Gu Z, and Shen K, Accuracy of estrogen receptor, progesterone receptor, and HER2 status between core needle and open excision biopsy in breast cancer: a meta-analysis. *Breast cancer research and treatment*, 2012 134(3): p. 957–967. [PubMed: 22370627]
37. Bueno-de-Mesquita JM, Nuyten DSA, Wesseling J, et al., The impact of inter-observer variation in pathological assessment of node-negative breast cancer on clinical risk assessment and patient selection for adjuvant systemic treatment. *Annals of Oncology*, 2010 21(1): p. 40–47. [PubMed: 19622588]
38. Reisenbichler ES, Lester SC, Richardson AL, et al., Interobserver Concordance in Implementing the 2010 ASCO/CAP Recommendations for Reporting ER in Breast Carcinomas A Demonstration of the Difficulties of Consistently Reporting Low Levels of ER Expression by Manual Quantification. *American Journal of Clinical Pathology*, 2013 140(4): p. 487–494. [PubMed: 24045544]

Figure 1:
Flow Diagram for selection of patient population.

Figure 2:

The lower triangular portion of the figure contains pie charts of the Pearson's correlation coefficient between any pair of measures after binarizing them. Completely filled pie chart indicates a correlation of 1 (when blue) or -1 (when orange). Partially filled pie charts, the direction of fill indicates the sign (positive for clockwise and negative for correlation) of correlation and the proportion indicates the values. Additionally, colors of the pie chart are bluish for high positive correlation and orangish for low correlation (positive or negative).

Figure 3:
Hazard ratios with confidence intervals for 31 measures of heterogeneity for the entire available population and subsets of patients that received specific therapies (detailed definitions of these measures are available in Supplementary Material eTable 1)

Figure 4:

Hazard Ratio for the entire available population and subsets of patients that received specific therapies using maximum difference of binarized ER Percentage Staining between samples as the measure of heterogeneity.

Table 1

List of abbreviations used in our study

Abbreviation	Full form
ER	Estrogen Receptor
PR	Progesterone Receptor
HER2	Human Epidermal Growth Factor Receptor2
EGFR	Epidermal Growth Factor Receptor
IHC	Immunohistochemical
DRFS	Distant Recurrence-Free Survival
ACIS	Automated Cellular Imaging System
FISH ratio	Fluorescence In Situ Hybridization ratio
HR	Hazard Ratio
MRI	Magnetic Resonance Imaging
SD	Standard Deviation
MDBS	Maximum Difference Between Samples using

Table 2

Clinicopathological characteristics of the entire study cohort and therapy specific cohort

Characteristics		All Patients (N* = 971)	Chemotherapy Patients (N* = 655)	Endocrine therapy Patients (N* = 680)	Anti-HER2 therapy Patients (N* = 168)
Age					
	Median (years)	52	50	52	47
	Range (years)	21–90	21–81	27–90	23–80
Race					
	White	705	446	516	114
	Black	201	161	118	39
	Asian	15	12	14	2
	Native	3	3	2	0
	Hispanic	18	13	10	7
	Multi	6	5	6	4
	American Indian	6	6	3	2
	Not Available (NA)	17	9	11	0
Menopausal Status					
	Pre	435	341	311	88
	Post	527	308	363	78
	NA	9	6	6	2
Tumor Size (based on TNM Staging)					
	T1	440	223	319	56
	T2	399	308	273	82
	T3	103	98	70	24
	T4	23	22	14	5
	NA	6	4	4	1
Median Follow up (months)		50	50	51	52
Number of distant Recurrences Multifocal/Multicentric?		85	80	48	16
	Yes	397	295	280	81
	No	573	359	399	86
	NA	1	1	1	1
ER Status					
	Positive	736	442	661	112
	Negative	211	193	14	46
	NA	24	20	5	10
PR Status					
	Positive	673	397	599	98
	Negative	275	239	77	61
	NA	23	19	4	9
HER2 Status					
	Positive	165	148	104	150

Characteristics	All Patients (N* = 971)	Chemotherapy Patients (N* = 655)	Endocrine therapy Patients (N* = 680)	Anti-HER2 therapy Patients (N* = 168)
Negative	787	493	567	10
NA	19	14	9	8

* indicates that this is the maximum number of patients that we have in the available population and for each therapy. The available population for each heterogeneity measure is less than this maximum value for available population and for each therapy.

Table 3

Mean correlation between heterogeneity measures belonging to same group and different groups

	Mean correlation of measures within same group	Mean correlation of measures with different groups
ER IHC	0.32	0.06
PR IHC	0.36	0.03
HER2 IHC	0.12	0.05
Proliferation (Ki-67)	0.42	0.03
EGFR IHC	0.52	0.03
Grade	0.31	0.04
Type	-	0.03