ARTICLE

HER2 in situ hybridization test in breast cancer: quantifying margins of error and genetic heterogeneity

António Polónia [1,2] · Ana Caramelo[1,2]

Abstract

The aim of the present study was to evaluate the effect of counting increasing number of invasive cancer cells in the result of the HER2 in situ hybridization (ISH) test in breast cancer as well as to compare two different approaches of measuring genomic heterogeneity (single cell and population based). A cohort of 100 consecutive breast cancer cases (primary and metastatic) were evaluated for HER2 gene amplification with bright-field ISH. The evaluation of the samples included scoring 20 nuclei, in five different areas, measuring the margins of error for each case. Genomic heterogeneity (GH) was defined by the 2018 ASCO/CAP guideline as a discrete population of tumor cells with HER2 amplification. We also evaluated GH as single tumor cells with HER2 amplification. The stabilization of the coefficient of variation of HER2/CEP17 ratio requires about 60 invasive cancer cells. The average margin of error of HER2/CEP17 ratio and of HER2 copy number was 0.40 and 0.53, respectively, when counting 20 cells, decreasing to 0.20 and 0.26 when counting 100 cells. Population GH was observed in 1% of the cases, while single cell GH was observed in 27% of the cases, reaching its maximum value in cases near the thresholds of positivity. Therefore, margins of error in HER2 ISH test are high, and the minimal cell number recommended in current guidelines should be raised to at least 60 cells. Population GH is a rare event and single cell GH is maximal in cases near the thresholds.

Introduction

Human epidermal growth factor receptor 2 (HER2) status can be determined by immunohistochemistry (IHC) or in situ hybridization (ISH) assays for the evaluation of HER2 overexpression or gene amplification, respectively. Usually, the assessment begins with IHC, with equivocal results requiring ISH reflex test. The 2018 ASCO/CAP (American Society of Clinical Oncology/College of American Pathologists) guideline recommends counting at least

20 non-overlapping cells in two separate areas of invasive cancer in HER2 ISH test [1]. Although this is usually interpreted as counting a total of 40 cells (at least 20 cells per area), the supplementary data of the previous guideline explains that the minimum cell number is, in fact, a total of 20 cells in two separate areas of invasive cancer (at least 10 cells per area) [2].

HER2 testing variability remains an important issue, with the minimum number of cells needed to obtain an accurate result yet to be determined [3–6]. Moreover, genetic heterogeneity has been reported in almost all types of cancer, including BC, being responsible for discrepant HER2 results between IHC and ISH tests, as well as contributing to changes in disease progression and response to target therapy [7, 8]. Several clinical studies have demonstrated that anti-HER2 targeted therapy improves progression-free survival and overall survival only in patients with HER2-positive BC, which represents about 15% of all BC cases [9–15]. For this reason, the accurate assessment of HER2 status is crucial to identify patients who are most likely to benefit from this targeted therapy.

In this study we aim to evaluate the effect of counting an increasing number of invasive cancer cells in the result of

✉ António Polónia
antoniopolonia@yahoo.com

1 Department of Pathology, Ipatimup Diagnostics, Institute of Molecular Pathology and Immunology, University of Porto, Porto, Portugal

2 I3S – Instituto de Investigação e Inovação em Saúde, University of Porto, Porto, Portugal

ISH quantification (HER2/CEP17 ratio and average of HER2 copy number per cell) as well as to compare two different approaches of measuring genetic heterogeneity (single cell and population based).

Materials and methods

Case selection

A cohort of 100 consecutive BC cases (primary and metastatic) with an equivocal HER2 result by IHC (score of 2+) was retrieved from the archives of Ipatimup Diagnostics from April to August 2019. The cases included formalin-fixed paraffin-embedded needle core biopsies (NCB) and surgical excision specimens (SES) referred to our institution (national reference center for HER2 ISH) for an evaluation of HER2 amplification with bright-field ISH. HER2 test by IHC was performed by the sending institution and information regarding pre-analytical conditions as well as the antibody used were not available. This study has been performed in accordance with the national regulative law for the handling of biological specimens from tumor banks, being the samples exclusively available for research purposes in retrospective studies, as well as under the international Helsinki declaration. Ethical approval and informed consent were not required for this study.

Bright field in situ hybridization

ISH was performed on 3-µm-thick sections in one block of each case with dual-hapten, dual-color ISH. The dual-probe assay (VENTANA HER2 Dual ISH DNA Probe Cocktail Assay (catalog number 760-6072); Ventana Medical Systems, Inc., Tucson, AZ, USA), which is Food and Drug Administration-approved, contains a HER2 locus-specific probe (black signal) and a control probe specific for the centromere of chromosome 17 (centromere enumeration probe-CEP17, red signal) that allows detection of HER2 amplification by light microscopy. The entire procedure was carried out on an automated staining system (Ventana BenchMark XT Staining System; Ventana Medical Systems, Inc., Tucson, AZ, USA) according to the manufacturer's instructions. Appropriated positive and negative controls were used in every set of slides. Optimal staining consists of an absence of non-specific background staining, distinct nuclear morphology, and clear and specific signals within the nucleus.

ISH interpretation

The samples were quantified by a biomedical scientist (AC) with previous ISH training according to the 2018 ASCO/

CAP guideline for HER2 amplification in BC [1]. The training, six months prior to this study, consisted in the parallel evaluation of 100 ISH tests with the pathologist achieving a diagnostic result concordance higher than 95% (amplified and non-amplified status) [16].

Corresponding hematoxylin and eosin staining was used for the identification of the invasive component of the tumor, and, whenever available, the IHC slide was used to score in the area with strongest intensity. The evaluation of the HE and the IHC slides were performed by a pathologist (AP) that marked the tumor areas (manually) for ISH quantification. Only cells with a minimum of one copy of HER2 and CEP17 each were scored. The number of HER2 signals was estimated in clusters, except for doublets, which counted as a single signal. The evaluation of the samples included scanning the entire ISH slide prior to counting and scoring 20 nuclei, in five different areas, recording the numbers of HER2 and CEP17 signals in each cell over an area with higher level of HER2 amplification. This approach allowed us to continuous add the result of individual cells, until 100 cells were reached, and measure the effect on the HER2/CEP17 ratio and on the average of HER2 copy number quantifications.

The 2018 BC guideline defines HER2 gene amplification as positive (classical group 1) when the HER2/CEP17 ratio is ≥2.0 and the average HER2 copy number is ≥4.0 signals per cell, and negative (classical group 5) when the HER2/CEP17 ratio is <2.0 and the average HER2 copy number is <4.0 signals per cell. Moreover, group 2 is defined as HER2/CEP17 ratio ≥2.0 and average HER2 copy number <4.0 signals per cell; group 3 as HER2/CEP17 ratio <2.0 and average HER2 copy number ≥6.0 signals per cell; and group 4 as HER2/CEP17 ratio <2.0 and average HER2 copy number ≥4.0 and <6.0 signals per cell. The final classification in groups 2 to 4 (non-classical groups) depends on the result of IHC analysis and is considered positive if a score 3+ in these groups or a score 2+ in group 3, and negative if otherwise [1].

The results obtained by this approach were compared with the results from the original ISH report performed by a pathologist (AP) counting 20 cells in two separate areas of invasive cancer (20 + 20), with an additional 20 cells if HER2/CEP17 ratio between 1.8 and 2.2. Cases with discordant results were reviewed during common microscopy session to search for reasons of disagreement. No additional testing was performed to resolve the discordances. Finally, genetic heterogeneity (GH) was documented, defined in the 2018 ASCO/CAP guideline as a discrete population of tumor cells with HER2 amplification. A case was considered positive if HER2 gene amplification represented at least 10% of the total tumor cell population [1]. We also evaluated single cell genetic heterogeneity, defined as tumor cells with HER2 amplification between 10 and 90% of the

total tumor cells without forming a discrete population of tumor cells.

Statistical analysis

Statistical analyses were performed using the Statistical Package for the Social Sciences (SPSS) version 26.0 for Windows. Pearson´s correlation coefficient (PCC) was used for comparison of quantitative variables. The level of significance was set at $P < 0.05$.

Interobserver agreement rates regarding interpretation of the HER2 amplification assay were evaluated with kappa (k) statistics. k-Values range between zero (chance agreement) and 1 (perfect agreement) and were satisfactory if greater than 0.81.

The coefficient of variation (CV) was used to quantify the HER2/CEP17 ratio and HER2 copy number variability independently of the unit of measurement, which results from the standard deviation divided by the mean. Margin of error (ME) at 95% confidence interval (CI) was calculated by multiplying the critical value (1.96) with the standard of error (SE). Standard of error was calculated as the ratio of standard deviation with the squared root of the number of cells analyzed. Curve estimation regression models were used to describe the behavior of the margins of error with an increasing number of invasive cancer cells.

Results

The cohort included 83 needle core biopsies and 17 surgical specimens, diagnosed in 98 women and in 2 men. The age of the patients ranged from 33 to 93 years old, with a median age at diagnosis of 60 years old. The histologic types of primary BC cases were invasive carcinoma of no special type (87), lobular carcinoma (7), and micropapillary carcinoma (2). The metastatic sites included regional lymph nodes (2), bone (1), and liver (1). Cohort characteristics are summarized in Table 1.

Regarding ISH quantification measurements, we observed a high correlation of HER2/CEP17 ratio and of average of HER2 copy number between the observers (PCC = 0.959; $p < 0.001$ and PCC = 0.916; $p < 0.001$, respectively) (Fig. S1A, B). Importantly, we observed a low correlation of HER2/CEP17 ratio and of average of HER2 copy number with the coefficient of variation (CV) (PCC = 0.284; $p = 0.004$ and PCC = 0.465; $p < 0.001$, respectively) (Fig. S2A, B), showing that this measure of variability is independent from the unit of measurement. We also observed that a minimum of 56 or 30 invasive cancer cells per case are required to stabilize the CV for HER2/CEP17 ratio or for average of HER2 copy number (difference from the final CV less than 0.01), respectively (Figs. 1A and

Table 1 Cohort characteristics.

	n
Sample	
(needle core biopsy/surgical specimen)	83/17
Gender	
(female/male)	98/2
Age	
(mean±sd)	60.76 ± 15.58
Primary	
Invasive carcinoma, NST	87
Lobular carcinoma	7
Micropapillary carcinoma	2
Metastases	
Regional lymph node	2
Bone	1
Liver	1

sd standard deviation, *NST* no special type.

S3A). The average margin of error of HER2/CEP17 ratio was determined with increasing number of cells, reaching a value of 0.40, 0.29, 0.25, 0.22 and 0.20 when counting 20, 40, 60, 80, and 100 invasive cancer cells, respectively (Fig. 1B). In addition, the average margin of error of average of HER2 copy number was also determined with increasing number of cells, reaching a value of 0.53, 0.41, 0.34, 0.29, and 0.26 when counting 20, 40, 60, 80, and 100 invasive cancer cells, respectively (Fig. S3B). Curve estimation regression models showed that a minimum of 457 or 926 invasive cancer cells per case are needed to reach a margin of error below 0.1 for HER2/CEP17 ratio (power model ($y = 1.48x^{-0.44}$), $R^2 = 0.998$, $p < 0.001$) or for average of HER2 copy number (power model ($y = 2.02x^{-0.44}$), $R^2 = 0.997$, $p < 0.001$), respectively.

In ISH test, HER2 positivity was identified in 15 cases by the pathologist (13 cases group 1 and 2 cases group 3) and in 14 cases by the biomedical scientist (12 cases group 1 and 2 cases group 3), with 97 concordant cases ($k = 0.879$). The discordances (positive versus negative) were observed between group 1 and groups 5 and 4 (2 and 1 cases, respectively). Nevertheless, there were 4 more cases with discordant negative ISH groups ($k = 0.786$). These cases corresponded to discordances between group 5 and groups 2 and 4 (3 and 1 cases, respectively), all considered negative given the equivocal result by IHC (for details see Tables 2 and 3).

Regarding the 3 discordant cases (positive/negative), only one case (case #17) had a HER2/CEP17 ratio and an average of HER2 copy number quantification with a margin of error that did not cross the threshold of different ISH groups. The remaining two cases, as well as the cases with discordances between negative ISH groups, had
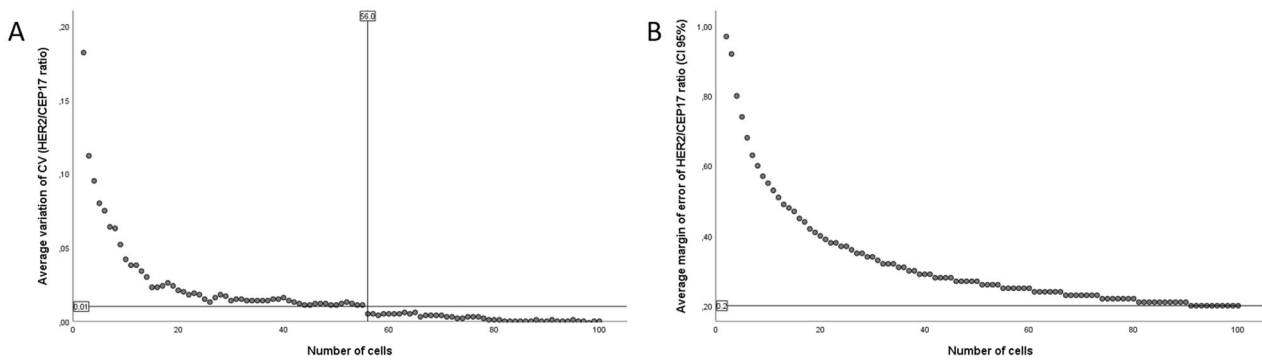
Fig. 1 Relationship between number of cells and the value of CV and ME of HER2/CEP17 ratio. A Average difference of the coefficient of variation (CV) of HER2/CEP17 ratio according to the number of cells and the final CV; **B** Average margin of error of HER2/CEP17 ratio (confidence interval of 95%) according to the number of cells.

Table 2 Distribution of cases in each ISH group.

		Biomedical scientist					
Pathologist	Group	1	2	3	4	5	Total
	1	11	0	0	0	2	13
	2	0	2	0	0	1	3
	3	0	0	2	0	0	2
	4	1	0	0	0	0	1
	5	0	2	0	1	78	81
	Total	12	4	2	1	81	100

HER2/CEP17 ratio or average HER2 copy number quantifications with margins of error that crossed the thresholds of different ISH groups. After common microscopy session, we detected the following reasons for discordant results: weak black signals (case #8), slight non-specific precipitation of black signals (case #17) and quantification of signals in non-invasive carcinoma (in situ carcinoma) (case #80) (Fig. 2). GH assessed by visual observation according to the ASCO/CAP 2018 guideline (discrete population of tumor cells with HER2 amplification) was not documented in any of the cases.

To objectively measure GH of different populations of tumor cells, we compared the HER2/CEP17 ratio and the average of HER2 copy number of different areas in each case, analyzing the overlap of margins of error. We observed that in 20% and 39% of the cases, there was at least 1 area with HER2/CEP17 ratio or average of HER2 copy number, respectively, different from the remaining areas. When using a combined threshold of HER2/CEP17 ratio of 2.0 and average of HER2 copy number of 4.0 and 6.0 (corresponding to different ISH groups), GH in different areas was observed only in 7% of the cases (Table 4). In these cases, 3 cases had positive and negative ISH areas, with only one case (case #70) showing one different area with a margin of error that did not cross the combined threshold. This case was classified as positive (ISH group

3), which was concordant with the original report, although the negative area (ISH group 4) was initially overlooked (Fig. 3A). The other two cases (cases #34 and #66) had a slight non-specific precipitation of black signals which caused the classification of just one area as positive, in each case, with average of HER2 copy number margins of error crossing the thresholds of negativity (Fig. 3B, C). Regarding the remaining 4 cases with different negative ISH areas, 3 cases had a combination of groups 5 and 2, and 1 case had a combination of groups 5 and 4. In addition, 2 of these heterogeneous cases (case #64 and #66) were part of the 7 cases with discordances between the observers.

Single cell GH (10% <positive cells <90%) was observed in 84, 34 and 8% of the cases when using the threshold of HER2/CEP17 ratio >2.0, and of HER2 copy number >4.0 or >6.0, respectively. However, when using a combined threshold of HER2/CEP17 ratio >2.0 and HER2 copy number >4.0, or HER2 copy number >6.0 regardless of ratio (corresponding to ISH groups 1 and 3, respectively), single cell GH was observed in 27% of the cases. We observed a high correlation between the number of positive cells observed in the cases and the expected number of positive cells if the cases had normal distributions of measurements (PCC = 0.973, $p < 0.001$; PCC = 0.988, $p < 0.001$ and PCC = 0.996, $p < 0.001$, using the threshold of HER2/CEP17 ratio of 2.0, and of average of HER2 copy number of 4.0 and 6.0, respectively) (Fig. S4). Finally, we evaluated the relationship of single cell GH with the values of HER2/CEP17 ratio and of average of HER2 copy number, observing that it reaches its maximum value near the thresholds of positivity (Figs. 4 and S5).

Discussion

In the last years, we have been documenting the effect of the latest ASCO/CAP guidelines for HER2 evaluation in BC [17, 18]. In the era of precision medicine, it becomes

Table 3 Cases with HER2 status discordance between observers.

Case	Gender	Age	Sample	IC	Pathologist			Biomedical scientist		
					HER2 status	HER2/CEP17 ratio	HER2 CN	HER2 status	HER2/CEP17 ratio	HER2 CN
8	F	51	NCB	NST	Positive (Group 1)	2.11	4.00	Negative (Group 5)	1.96	2.87
10	F	62	SES	NST	Negative (Group 5)	1.19	2.98	Negative (Group 4)	1.36	4.00
17	M	38	NCB	NST	**Negative (Group 4)**	**1.65**	**4.62**	**Positive (Group 1)**	**2.33**	**6.03**
64	F	57	NCB	NST	Negative (Group 5)	1.91	2.48	Negative (Group 2)	2.31	2.66
66	F	58	NCB	NST	Negative (Group 5)	1.85	3.10	Negative (Group 2)	2.26	3.16
80	F	34	NCB	NST	Positive (Group 1)	3.29	4.05	Negative (Group 5)	1.86	2.44
86	F	62	NCB	NST	Negative (Group 2)	2.25	2.70	Negative (Group 5)	1.91	2.24

Bold—case with ISH quantifications with margins of error not crossing the combined threshold.

F female, *M* male, *SES* surgical excision specimen, *NCB* needle core biopsy, *NST* no special type.
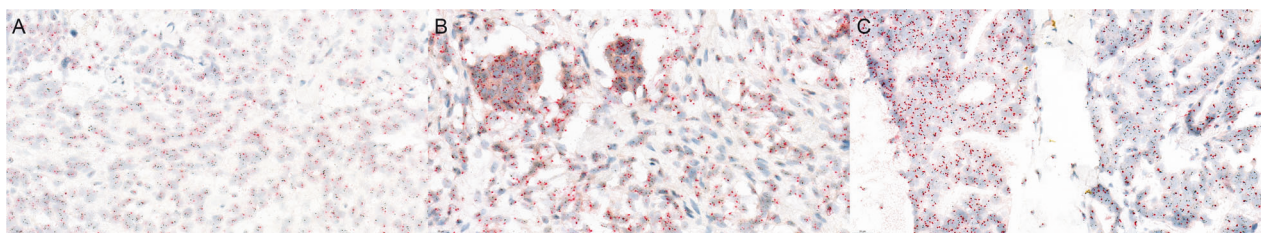

Fig. 2 ISH images from cases with HER2 status discordance between observers. **A**: case #8 with weak black signals; **B**: case #17 with slight non-specific precipitation of black signals; **C**: case #80 with invasive carcinoma (right) with HER2 amplification and ductal carcinoma in situ (left) without HER2 amplification.

important not only to provide the correct result but, whenever applicable, the most precise one. The ISH assay for HER2 in BC has the objective to quantify HER2 gene amplification, supplying a continuous result. The importance of precision becomes even more relevant in cases in which the result is near the decision thresholds. As far as we know, there are no studies evaluating the exact precision of this assay in the context of BC. In this work, we show that about 60 invasive cancer cells are required to stabilize the coefficient of variation of HER2/CEP17 ratio, a number that many laboratories already use empirically, and that we now show the underlying mathematical reasoning. In addition, the precision, when following the 2018 ASCO/CAP guidelines, is very low, having margins of error of 0.40 for HER2/CEP17 ratio and 0.53 for average of HER2 copy number, when counting 20 invasive cancer cells (the minimum number of cells required). Even when increasing the number of invasive cells to 100 (which is rarely done) the margins of error are not below 0.20. This means that in cases with HER2/CEP17 ratio between 1.8 and 2.2, when at least 40 cells should be evaluated, it is very likely that the final result has margins of error crossing the decision thresholds. As such, we continue to defend that 20 cells should not be the minimal cell number recommended by current guidelines [19]. Although the number of cases could be higher in this study, a limitation of this work, we propose

that laboratories should quantify their margins of error, including it in the HER2 ISH report.

The 2018 ASCO/CAP guideline already mentions the problem of cases with HER2 evaluations near decision thresholds, although the explanation provided by the authors, in our opinion, may be an unfortunate one as it states that "there is a high likelihood that repeat testing will result in different results by chance alone" [1]. Instead of giving the impression that results can be random, it should be point out that any quantification measurement in any field can have imprecision and HER2 gene amplification is not an exception to this rule. We have shown previously that intra and interobserver concordance rate of HER2 ISH test increases with increasing cell count from 20 to 60 invasive cells [19]. We have also shown, along with others, that cases near the thresholds are precisely the cases that can have intra-observer as well as interobserver discordances, making a precise result even more clinically relevant [19, 20]. Our data estimates that about 450 or 930 invasive cells must be evaluated to reach margins of error of 0.1 (for HER2/CEP17 ratio and average of HER2 copy number, respectively), which is impractical for manual assessment. To overcome this limitation, image analysis of ISH tests can assist this quantification having the potential to evaluate thousands of cells, lowering the margins of error to minimal values. Clinicians should be aware of the limitations of

Table 4 Cases with genetic heterogeneity in different areas.

Case	Gender	Age	Sample	IC	Area 1	2	3	4	5
34	F	81	NCB	NST	R 1.97 CN 3.25 ISH G5	R 1.77 CN 3.60 ISH G5	R 1.34 CN 2.90 ISH G5	R 1.35 CN 3.15 ISH G5	R 2.42 CN 4.05 ISH G1
64	F	57	NCB	NST	R 2.38 CN 2.85 ISH G2	R 3.10 CN 3.25 ISH G2	R 2.06 CN 2.55 ISH G2	R 1.85 CN 2.05 ISH G5	R 2.18 CN 2.06 ISH G2
66	F	58	NCB	NST	R 1.94 CN 3.05 ISH G5	R 1.98 CN 2.70 ISH G5	R 2.88 CN 4.10 ISH G1	R 2.05 CN 3.05 ISH G2	R 2.45 CN 2.90 ISH G2
70	F	66	NCB	NST	R 1.54 CN 6.75 ISH G3	**R 1.14 CN 4.55 ISH G4**	R 2.68 CN 8.30 ISH G1	R 2.01 CN 8.00 ISH G1	R 1.75 CN 5.85 ISH G4
75	F	84	NCB	NST	R 1.14 CN 2.55 ISH G5	R 1.64 CN 3.45 ISH G5	R 1.42 CN 4.55 ISH G4	R 1.23 CN 2.75 ISH G5	R 1.44 CN 3.80 ISH G5
90	F	37	NCB	NST	R 1.70 CN 1.75 ISH G5	R 1.30 CN 1.85 ISH G5	R 2.43 CN 2.90 ISH G2	R 1.42 CN 1.85 ISH G5	R 1.80 CN 2.40 ISH G5
100	F	40	NCB	NST	R 1.76 CN 2.30 ISH G5	R 1.83 CN 1.95 ISH G2	R 2.25 CN 2.35 ISH G2	R 2.50 CN 2.70 ISH G2	R 2.30 CN 2.75 ISH G2

Bold—different ISH group (positive/negative) quantification with margins of error not crossing the combined threshold.

F female, NCB needle core biopsy, NST no special type.

current measurements, integrating the results, preferably with margins of error, along with the traditional criteria to decide the best treatment for individual patients.

In our work, we observed a high correlation of measurements between both observers and a concordance rate above 95%, with only one discordant case with margins of error that did not cross the decision thresholds. Although the invasive cancer was marked on HE by the pathologist, we were able to trace a case in which the quantification was made on in situ carcinoma (mixed with invasive carcinoma) underlying the importance of training, experience and supervision in the interpretation of HER2 ISH test. The UK guidelines recommend that when training new professionals in HER2 ISH test, evaluations of at least 100 ISH tests in parallel with an experienced observer should be done until a minimum concordance of 95% is reached [16].

The first definition of HER2 genetic heterogeneity (HER2-GH) was published in 2009 as an extension of the 2007 ASCO/CAP guideline described as HER2 gene amplification in 5 to 50% of individual invasive cancer cells [8]. Afterwards, several studies reported the presence of HER2-GH in BC from 5 to 40% of the cases [21–26]. The definition was not based on clinical studies with prognosis or response to target therapy, representing the first step to unravel the clinical significance of HER2-GH. Later, it was shown that the presence of HER2-GH was associated with reduced disease-free survival and less response to anti-HER2 target therapy [25, 26].

In this study, we show that single cell HER2-GH can be accurately estimated from the values of HER2/CEP17 ratio and average of HER2 copy number, being higher near the thresholds of positivity. Previously, it was demonstrated that HER2-GH is more frequent in cases with low HER2 amplification, ratios near the thresholds and that GH measured in individual cells is not informative of clonal heterogeneity within a tumor population [24, 25, 27]. The different prognosis, as well as response to target therapy, in BC cases with single cell HER2-GH can be explained by the closeness of HER2/CEP17 ratio and HER2 copy number values to the thresholds, representing nothing more than a mathematical artifact, instead of representing the presence of different biological clones.

We acknowledge that the definition of single cell HER2-GH used in this study is different from the initial definition proposed in 2009. The upper limit of 50% was soon criticized arguing that if HER2-GH is defined as the presence of more than 5% of amplified cells, the 95% should be used as the upper limit. The application of the original definition could give rise to paradox cases, with difficult interpretation and unknown clinical significance [28, 29]. In this study, we choose to use the same proportion of positivity in both definitions (10–90%) to compare the effect of measuring HER2-GH using single cells or population of cells,
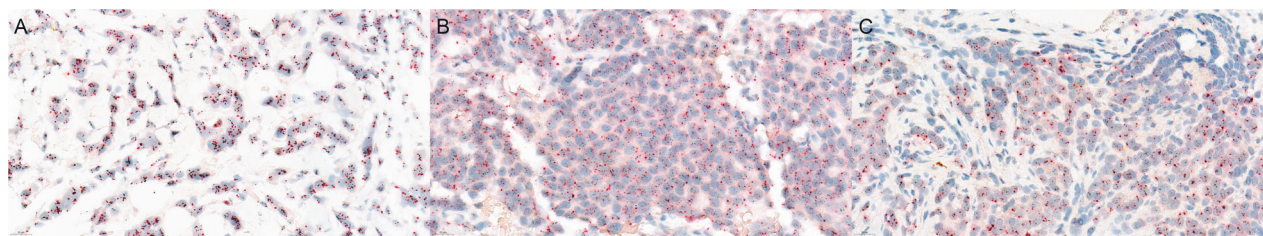
Fig. 3 ISH images from cases with genetic heterogeneity in different areas. A: case #70 ISH group 3; **B** and **C**: case #34 and #66 with slight non-specific precipitation of black signals.
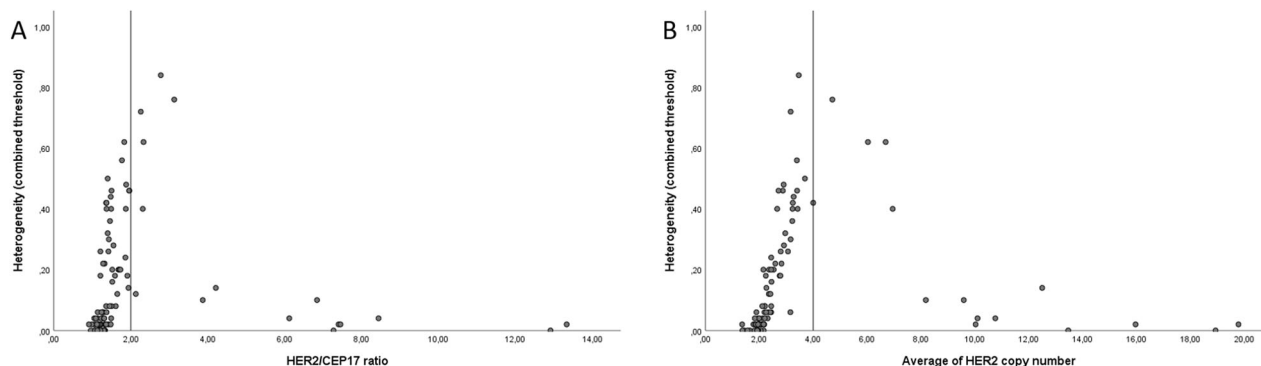

Fig. 4 Relationship between heterogeneity and the value of HER2/CEP17 ratio and HER2 copy number. A relationship between the HER2/CEP17 ratio and the proportion of single cell genetic heterogeneity using the combined threshold; **B**: relationship between the HER2 copy number and the proportion of single cell genetic heterogeneity using the combined threshold.

excluding other differences as a source of discrepancy in the results.

The 2013 ASCO/CAP changed the definition of HER2-GH from individual cells to discrete population of tumor cells (at least 10% of the total tumor cell population with HER2 amplification) [2]. After this change, the presence of HER2-GH has been reported in less than 2% of the cases, although there is still the need to determine the minimal proportion of amplified tumor cell population that achieves clinical response to HER2-targeted therapy [17, 18, 30, 31]. Interestingly, it was shown that in BC cases with HER2-GH up to 30%, both cases with clustered and scattered amplified cells showed an intermediate clinical behavior between homogenous amplified and non-amplified BC [31]. Moreover, HER2-GH by IHC (score 3+ in >10 and <100%) was not considered in this study, which focused rather on ISH analysis in equivocal results (score of 2+). Importantly, HER2-GH can be missed when just evaluating core biopsies rather than excision specimens, although it is less frequently observed than in other breast biomarkers (estrogen and progesterone receptors) [32].

In this work, we clearly show that single cell HER2-GH is much higher than population HER2-GH in equivocal (score of 2+) BC cases. When using HER2/CEP17 ratio higher than 2.0 or average of HER2 copy number higher than 4.0, we can document GH in about 85% or a third of the cases, respectively. Even when using a combined

threshold, corresponding to different ISH groups, we still report single cell HER2-GH in about 30% of the cases, which is much higher than the population HER2-GH observed (less than 10% of the cases). In these last cases, only 3 cases had a mixture of positive and negative areas, and only one case (case #70) had quantifications with margins of error far from the thresholds, as well as without artifacts, supporting true HER2-GH. Although we could not find the negative area in that case after review, the quantification of the average of HER2 copy number above 4.0 makes it unlikely that the measurement was made in an area other than invasive carcinoma, probably representing a minor component. Nonetheless, the impact of the presence of minor components (either HER2 amplified or not) needs to be further studied and clarified. Unfortunately, the absence of clinical data in this cohort (such as survival data or response to target-therapy) made it impossible to correlate it with HER2-GH.

Cases with population HER2-GH as defined by different negative ISH groups (5, 4, and 2) represent heterogenous negative BC cases that might display different clinical behavior compared to homogenous negative group 5 BC cases, a situation that, although rare in this study (less than 5%), has not been investigated in the literature. It has already been shown that BC cases from group 4 and 2 have worse clinical outcome compared to group 5 [33, 34]. Importantly, most BC cases (more than 90%) disclosed homogeneity in HER2/CEP17 ratio and in average of HER2

copy number quantifications, with only 1% of the cases showing convincing population HER2-GH.

In conclusion, we show that margins of error in HER2 ISH test are high, even when counting 100 cells, a limitation that could be overcome through quantification with image analysis. In addition, we show that population HER2-GH is a rare event, and that single cell HER2-GH is maximal in cases near the thresholds of positivity.

Compliance with ethical standards

References

1. 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. Arch Pathol Lab Med. 2018;142:1364–82.
2. Wolff AC, Hammond ME, Hicks DG, Dowsett M, McShane LM, Allison KH, et al. Recommendations for human epidermal growth factor receptor 2 testing in breast cancer: American Society of Clinical Oncology/College of American Pathologists clinical practice guideline update. J Clin Oncol. 2013;31:3997–4013.
3. Yildiz-Aktas IZ, Dabbs DJ, Bhargava R. The effect of cold ischemic time on the immunohistochemical evaluation of estrogen receptor, progesterone receptor, and HER2 expression in invasive breast carcinoma. Mod Pathol. 2012;25:1098–105.
4. Moatamed NA, Nanjangud G, Pucci R, Lowe A, Shintaku IP, Shapourifar-Tehrani S, et al. Effect of ischemic time, fixation time, and fixative type on HER2/neu immunohistochemical and fluorescence in situ hybridization results in breast cancer. Am J Clin Pathol. 2011;136:754–61.
5. Carbone A, Botti G, Gloghini A, Simone G, Truini M, Curcio MP, et al. Delineation of HER2 gene status in breast carcinoma by silver in situ hybridization is reproducible among laboratories and pathologists. J Mol Diagn. 2008;10:527–36.
6. Dietel M, Ellis IO, Hofler H, Kreipe H, Moch H, Dankof A, et al. Comparison of automated silver enhanced in situ hybridisation (SISH) and fluorescence ISH (FISH) for the validation of HER2 gene status in breast carcinoma according to the guidelines of the American Society of Clinical Oncology and the College of American Pathologists. Virchows Arch. 2007;451:19–25.
7. Polyak K. Heterogeneity in breast cancer. J Clin Invest. 2011;121:3786–8.
8. Vance GH, Barry TS, Bloom KJ, Fitzgibbons PL, Hicks DG, Jenkins RB, et al. Genetic heterogeneity in HER2 testing in breast cancer: panel summary and guidelines. Arch Pathol Lab Med. 2009;133:611–2.
9. Perez EA, Suman VJ, Davidson NE, Martino S, Kaufman PA, Lingle WL, et al. HER2 testing by local, central, and reference laboratories in specimens from the North Central Cancer Treatment Group N9831 intergroup adjuvant trial. J Clin Oncol. 2006;24:3032–8.
10. Chen X, Yuan Y, Gu Z, 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 Res Treat. 2012;134:957–67.
11. Rakha EA, Starczynski J, Lee AH, Ellis IO. The updated ASCO/CAP guideline recommendations for HER2 testing in the management of invasive breast cancer: a critical review of their implications for routine practice. Histopathology. 2014;64:609–15.
12. Cobleigh MA, Vogel CL, Tripathy D, Robert NJ, Scholl S, Fehrenbacher L, et al. Multinational study of the efficacy and safety of humanized anti-HER2 monoclonal antibody in women who have HER2-overexpressing metastatic breast cancer that has progressed after chemotherapy for metastatic disease. J Clin Oncol. 1999;17:2639–48.
13. Slamon DJ, Leyland-Jones B, Shak S, Fuchs H, Paton V, Bajamonde A, et al. Use of chemotherapy plus a monoclonal antibody against HER2 for metastatic breast cancer that overexpresses HER2. N Engl J Med. 2001;344:783–92.
14. Piccart-Gebhart MJ, Procter M, Leyland-Jones B, Goldhirsch A, Untch M, Smith I, et al. Trastuzumab after adjuvant chemotherapy in HER2-positive breast cancer. N Engl J Med. 2005;353:1659–72.
15. Vogel CL, Cobleigh MA, Tripathy D, Gutheil JC, Harris LN, Fehrenbacher L, et al. Efficacy and safety of trastuzumab as a single agent in first-line treatment of HER2-overexpressing metastatic breast cancer. J Clin Oncol. 2002;20:719–26.
16. Rakha EA, Pinder SE, Bartlett JM, Ibrahim M, Starczynski J, Carder PJ, et al. Updated UK Recommendations for HER2 assessment in breast cancer. J Clin Pathol. 2015;68:93–99.
17. Polonia A, Leitao D, Schmitt F. Application of the 2013 ASCO/CAP guideline and the SISH technique for HER2 testing of breast cancer selects more patients for anti-HER2 treatment. Virchows Arch. 2016;468:417–23.
18. Curado M, Caramelo AS, Eloy C, Polonia A. What to expect from the 2018 ASCO/CAP HER2 guideline in the reflex in situ hybridization test of immunohistochemically equivocal 2+ cases? Virchows Arch. 2019;475:303–11.
19. Polonia A, Eloy C, Pinto J, Braga AC, Oliveira G, Schmitt F. Counting invasive breast cancer cells in the HER2 silver in-situ hybridization test: how many cells are enough? Histopathology. 2017;71:247–57.
20. Dowsett M, Hanna WM, Kockx M, Penault-Llorca F, Ruschoff J, Gutjahr T, et al. Standardization of HER2 testing: results of an international proficiency-testing ring study. Mod Pathol. 2007;20:584–91.
21. Allison KH, Dintzis SM, Schmidt RA. Frequency of HER2 heterogeneity by fluorescence in situ hybridization according to CAP expert panel recommendations: time for a new look at how to report heterogeneity. Am J Clin Pathol. 2011;136:864–71.
22. Hanna WM, Ruschoff J, Bilous M, Coudry RA, Dowsett M, Osamura RY, et al. HER2 in situ hybridization in breast cancer: clinical implications of polysomy 17 and genetic heterogeneity. Mod Pathol. 2013;27:4–18.

23. Tubbs RR, Hicks DG, Cook J, Downs-Kelly E, Pettay J, Hartke MB, et al. Fluorescence in situ hybridization (FISH) as primary methodology for the assessment of HER2 Status in adenocarcinoma of the breast: a single institution experience. Diagn Mol Pathol. 2007;16:207–10.

24. Chang MC, Malowany JI, Mazurkiewicz J, Wood M. "Genetic heterogeneity" in HER2/neu testing by fluorescence in situ hybridization: a study of 2,522 cases. Mod Pathol. 2012;25:683–8.

25. Seol H, Lee HJ, Choi Y, Lee HE, Kim YJ, Kim JH, et al. Intratumoral heterogeneity of HER2 gene amplification in breast cancer: its clinicopathological significance. Mod Pathol. 2012;25:938–48.

26. Bartlett AI, Starcyznski J, Robson T, Maclellan A, Campbell FM, van de Velde CJ, et al. Heterogeneous HER2 gene amplification: impact on patient outcome and a clinically relevant definition. Am J Clin Pathol. 2011;136:266–74.

27. Buckley NE, Forde C, McArt DG, Boyle DP, Mullan PB, James JA, et al. Quantification of HER2 heterogeneity in breast cancer-implications for identification of sub-dominant clones for personalised treatment. Sci Rep. 2016;6:23383.

28. Hsu CY, Li AF, Yang CF, Ho DM. Proposal of modification for the definition of genetic heterogeneity in HER2 testing in breast cancer. Arch Pathol Lab Med. 2010;134:163. 162; author reply

29. Albarracin C, Edgerton ME, Gilcrease MZ, Huo L, Krishnamurthy S, Middleton LP, et al. Is it too soon to start reporting HER2 genetic heterogeneity? Arch Pathol Lab Med. 2010;134:162–3. author reply 163.

30. Polonia A, Oliveira G, Schmitt F. Characterization of HER2 gene amplification heterogeneity in invasive and in situ breast cancer using bright-field in situ hybridization. Virchows Arch. 2017;471:589–98.

31. Di Oto E, Brandes AA, Cucchi MC, Foschini MP. Prognostic impact of HER-2 subclonal amplification in breast cancer. Virchows Arch. 2017;471:313–9.

32. VandenBussche CJ, Cimino-Mathews A, Park BH, Emens LA, Tsangaris TN, Argani P. Reflex estrogen receptor/progesterone receptor/human epidermal growth factor receptor 2 (ER/PR/Her2) analysis of breast cancers in needle core biopsy specimens dramatically increases health care costs. Am J Surg Pathol. 2015;39:939–47.

33. Biserni GB, Engstrom MJ, Bofin AM. HER2 gene copy number and breast cancer-specific survival. Histopathology. 2016;69:871–9.

34. Wang X, Teng X, Ding W, Sun K, Wang B. A clinicopathological study of 30 breast cancer cases with a HER2/CEP17 ratio of >/=2.0 but an average HER2 copy number of <4.0 signals per cell. Mod Pathol. 2020;33:1557–62.