Navigating Breast Health: From Screening to Survivorship - Research Article

WOMEN'S HEALTH

Women's Health
Volume 21: 1–10
© The Author(s) 2025
Article reuse guidelines:
sagepub.com/journals-permissions
DOI: 10.1177/17455057241304654
journals.sagepub.com/home/whe
Sage

Unlocking breast cancer in Brazilian public health system: Using tissue microarray for accurate immunohistochemical evaluation with limitations in subtyping

Rubia Denise Ruppenthal[1,2,3], Emily Ferreira Salles Pilar[4], Jordan Boeira dos Santos[3], Rafael Correa Coelho[5], Carina Machado Costamilan Henriques[2], Diego de Mendonça Uchôa[2] and Marcia Silveira Graudenz[1,2,5]

Abstract

Background: Breast cancer (BC) is a significant burden on healthcare systems, especially in low- and middle-income countries where access to diagnosis and treatment is challenging.

Objectives: The purpose of this study was to assess the diagnostic accuracy and cost using tissue microarray (TMA) instead of traditional immunohistochemical (IHC) evaluation for estrogen receptor (ER), progesterone receptor (PR), human epidermal growth factor 2 (HER2), and the proliferation marker Ki-67 and BC subtyping within the Brazilian public health system.

Design: This is a retrospective cohort study comparing TMA slides with traditional whole-slide evaluation for IHC markers in 242 BC cases.

Methods: We used formalin-fixed tissue blocks for TMA assembly. Clinical data and IHC scores for ER, PR, HER2, and Ki-67 were obtained from pathology reports. Cohen's kappa (k) was used to assess TMA performance.

Results: BC samples were distributed in 10 TMAs and 968 cores were scored (242 BC cases $\times$ 4 markers). In 97% of these, TMA reached high quality to adequate IHC scoring with minimal technical issues. Inter-examiner agreement was almost perfect for all markers (ranging from 0.85 for HER2 to 0.91 for ER, $p < 0.001$). The intratumoral heterogeneity ranged from almost perfect agreement for ER and HER2 to moderate to substantial for PR and Ki-67. TMA offers substantial time and cost savings, with an approximately 11-fold reduction compared to traditional methods. The concordance between TMA and original reports was almost perfect, with 93% overall agreement ($k = 0.81$, $p < 0.001$). However, TMA performance varied between markers, with intratumoral heterogeneity significantly impacting discordant results, particularly for Ki-67 and HER2. This ultimately affected the accuracy of BC subtyping. TMA performed well in identifying luminal A and triple-negative cases, but misclassification was common for luminal B and HER2-positive cases.

Conclusion: TMA offers accurate and lower-cost results in the individualized IHC assessment of BC markers. However, we do not recommend the use of TMA in the subtyping of BC, where analysis of the whole section remains necessary

[1]Department of Pathology, Faculdade de Medicina, Universidade Federal do Rio Grande do Sul, Porto Alegre, Rio Grande do Sul, Brazil
[2]Laboratory of Pathology, Hospital de Clínicas de Porto Alegre, Porto Alegre, Rio Grande do Sul, Brazil
[3]Graduate Program in Sciences in Gastroenterology and Hepatology, Faculdade de Medicina, Universidade Federal do Rio Grande do Sul, Porto Alegre, Rio Grande do Sul, Brazil
[4]Laboratory of Experimental Pathology, Hospital de Clínicas de Porto Alegre, Porto Alegre, Rio Grande do Sul, Brazil
[5]Graduate Program in Medical Science, Faculdade de Medicina, Universidade Federal do Rio Grande do Sul, Porto Alegre, Rio Grande do Sul, Brazil

Corresponding author:
Rubia Denise Ruppenthal, 1210-2929 Aquitaine Avenue, Mississauga, L5N 2C7, ON, Canada.
Emails: rubia.ruppenthal@ufrgs.br; rubia.ruppenthal@uwo.ca

for more accurate results. We advocate more studies using the TMA approach in the Brazilian public health system to advance women's health care.

Keywords
breast cancer, immunohistochemistry, tissue microarray, public health, subtyping

Date received: 23 May 2024; revised: 14 October 2024; accepted: 5 November 2024

Introduction

Breast cancer (BC) is the most common cancer in women worldwide, with 70% of deaths from the disease occurring in low- and middle-income countries, such as Brazil.[1] The National Cancer Institute estimates that approximately 18,000 Brazilian women die of BC every year.[2] Around 75% of the population has no private health insurance and relies exclusively on the Universal Health System (SUS),[3] the largest public health system in the world that provides free healthcare to all Brazilians, regardless of their socioeconomic status. BC is more frequently diagnosed in its symptomatic and in more advanced stages in SUS than in private health systems[4] or high-income countries.[5]

Brazilian public hospitals face enormous pressure to optimize healthcare services and reduce costs.[6] The Hospital de Clínicas de Porto Alegre, a tertiary public hospital in the South of Brazil, processes approximately 540 immunohistochemical (IHC) tests of BC biomarkers per year at a cost of around 31,000 USD (154,400.00 BRL—Brazilian reais). This is the most common and expensive individual test offered in our laboratory and includes analysis of the expression of estrogen receptor (ER), progesterone receptor (PR), human epidermal growth factor 2 (HER2), and the proliferation marker Ki-67. These biomarkers are combined for BC subtyping into luminal (A and B), HER2-positive, and triple-negative tumors and guide systemic therapy. Proposing strategies to increase access to BC diagnosis and treatment is a priority in the Brazilian public health context.

The IHC tests of BC biomarkers are traditionally done on surgical specimens or biopsies on whole individual glass slides. The tissue microarray (TMA) approach, which combines multiple cylindrical fragments of tumor tissue from different patients in the same glass slide, has been extensively used in pathology research.[7] TMA saves working time, standardizes reactions, allows for comparative interpretation of cases, and reduces the total cost of tissue analyses.[8] However, the use of TMA in clinical practice remains controversial worldwide, and its feasibility and cost-benefit have never been evaluated in the Brazilian public health system before. BC was chosen as the prototype for this type of study due to its high regional prevalence at the regional level and throughout the country.

This study aimed to assess the diagnostic accuracy of TMA as a cost-effective alternative to evaluating the IHC status of ER, PR, HER2, Ki-67, and BC subtyping and maximize its potential use in clinical practice.

Methods

Patients

The study is a retrospective cohort analysis that evaluates the diagnostic accuracy of TMA in BC IHC evaluation. Two hundred forty-two women diagnosed with invasive BC in Hospital de Clínicas de Porto Alegre between 2010 and 2015 were consecutively included in the study. The patient eligible criteria are BC diagnosis and previous IHC evaluation for ER, PR, HER2, and Ki-67 available in medical records. Formalin-fixed tissue blocks from all patients were retrieved from the Laboratory of Pathology archive in accordance with ethical guidelines. We consistently follow the established preanalytical handling guidelines of the College of American Pathologists.

The clinical data and the original IHC scores of ER, PR, HER2, and Ki-67 were obtained from the anatomopathological reports through analyses of the whole slide and medical records. The average age was 58.2 years (range 24–92 years), and invasive carcinoma of the non-special type was the most frequent histopathological type of tumor. Pathological staging was determined using the AJCC TNM System,[9] and was distributed as follows: 131 patients in stage I, 57 in stage II, 42 in stage III, and 12 in stage IV. In 237 of the 242 cases, the IHC scores were fully available in the pathology report, making it possible to define the IHC subtype (BC subtype): 101 tumors were classified as luminal A (ER^+ and/or PR^+, $HER2^-$ and Ki-67 $\leqslant$20%), 87 as luminal B (ER^+ and/or PR^+, $HER2^+$ or Ki-67 $>$20%), 19 as HER2 positive (ER^-, PR^- and $HER2^+$) and 30 as triple negative (ER^-, PR^-, and $HER2^-$). Cases with tumor areas smaller than 2 cm, treated with neoadjuvant chemotherapy prior to surgical resection or without IHC evaluation for the four markers evaluated were excluded. Only excisional samples were utilized, as core biopsies were not employed to ensure the preservation of the patient's archived tissue and to mitigate the risk of material depletion during TMA assembly.

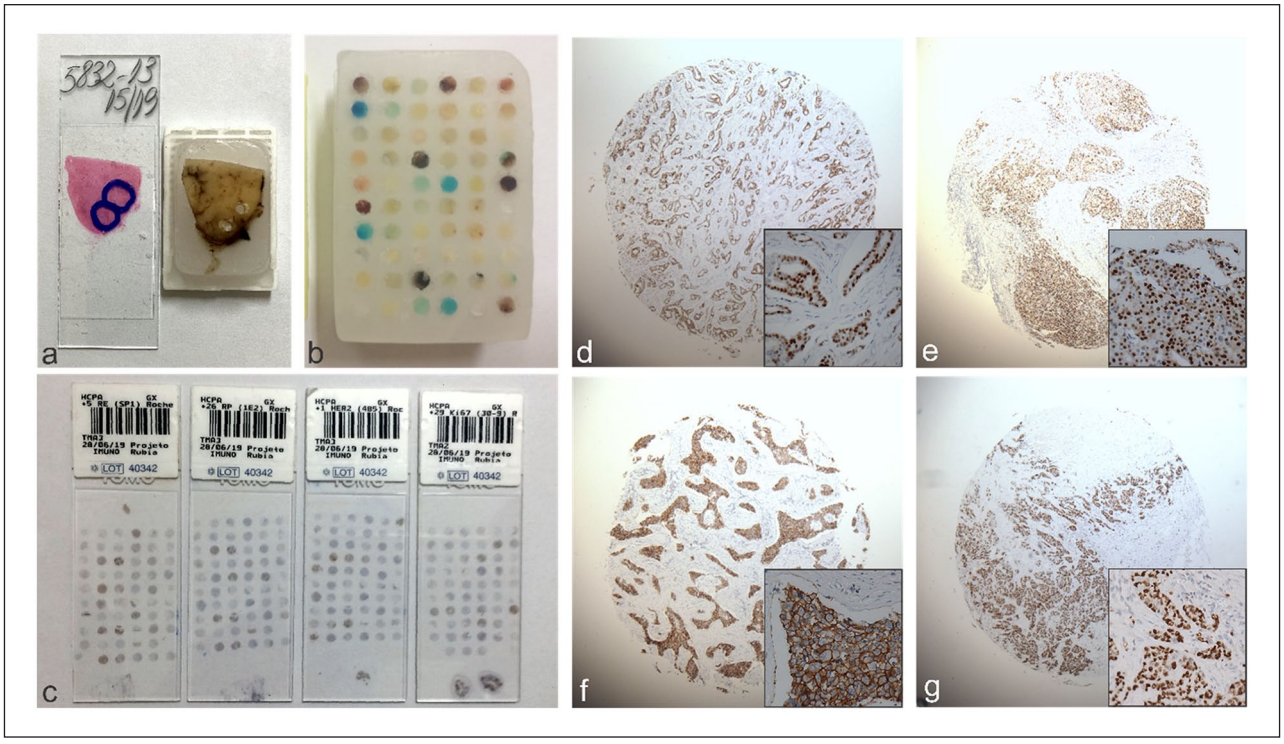

Figure 1. TMA construction and immunohistochemical analysis. (a) Selection of the most representative area on the H&E slide for extraction of the tumor core; (b) complete TMA block; (c) immunohistochemical slides for ER, PR, HER2, and Ki-67 of TMA sections and respective controls. Examples of cases classified as positive for ER (d), and PR (e). HER2 (f) and high Ki-67 (g), 100×. In the detail squares, the same cases at 400× magnification.
TMA: tissue microarray; ER: estrogen receptor; PR: progesterone receptor; HER2: human epidermal growth factor 2.

TMA assembly and immunohistochemistry

The most representative area of the tumor was carefully circled by an expert breast pathologist (MSG) on the hematoxylin-eosin-stained slide (Figure 1(a)) in areas with high tumor cellularity. For TMAs assembly, we used the manual TMA T-Sue system (Simport® Scientific, Beloeil, Canada) to extract two cores of 2.0 mm of each tumor using the principles first described by Kononen et al.[7] (Figure 1(b)). Briefly, the procedures began with the preparation of the TMA grid to correctly identify the position of each sample and the organization of the donor blocks. Then, two cylindrical tissue cores were extracted from the donor block with a 2.00 mm punch needle, no more than 3 mm deep, and precisely placed in the recipient block, which was previously prepared using the M473-60 mold. This mold has a capacity for 60 cores, distributed over 6 rows and 10 columns, allowing to include 24 duplicate tumors/cases per TMA. For guidance in reading the TMA, a core containing placental tissue was included in each TMA block. The cores were fixed with light pressure followed by brief heating, cooled overnight, and sectioned (4 µm). The sections were then mounted on slides for H&E staining and analysis. The total tumor area size analyzed is 3.14 mm^2 for each 2.0 mm core. The minimum number of tumor cells sufficient for scoring was ⩾100 per core.

For immunohistochemistry, the TMA blocks were cut into 3 µm sections and placed on glass slides with positive and negative controls. The sections were processed on Ventana automation equipment (BenchMark AutoStainer; Ventana Medical Systems, Tucson, AZ, USA) using the following antibodies: ER (clone SP1; Ventana, Tucson, AZ, USA), PR (clone 1E2; Ventana Medical Systems), HER2 (clone 4B5; Ventana Medical Systems) and Ki-67 (clone 30-9; Ventana Medical Systems) (Figure 1(c)). This immunostaining method is the same one used in the laboratory's routine work, with quality attested to by the Joint Commission International Accreditation Seal in 2017.

Microscopic analysis of TMA

TMA consolidated multiple tissue samples into a single slide for simultaneous analysis. In contrast, in the traditional IHC, a whole section of the tumor tissue is analyzed in slides individually. The reliability of TMA microscopic analysis depended on the quality of the TMA, the alignment of the cores, and the pathologist's ability to orient themselves and identify precise samples according to the grid. Then, following an initial overall TMA quality evaluation and positioning, the pathologist proceeded to assign a core-specific IHC score to each tissue core by traversing the slide in an up-and-down motion.

The same criteria of immunostaining evaluation in the whole section were applied to TMA. The evaluation of ER, PR, and HER2 expression was carried out in accordance with the guidelines of the American Society of Clinical Oncology.[10] Nuclear staining was considered positive for ER and/or PR when detected in at least 1% of tumor cells at any intensity (Figure 1(d) and (e)). For HER2, staining in the membranes of tumor cells was classified as follows[11]: 0, when no tumor cells showed HER2-positive staining or incomplete and weakly perceptible membrane staining in ⩽10% of tumor cells; 1+, incomplete and weakly perceptible staining in ⩾10% of tumor cells; 2+, weak to moderate complete staining observed in ⩾10% of tumor cells; and 3+, circumferential and strong complete staining in ⩾10% of tumor cells. Cases with a score of 3+ were considered HER2-positive. Cases 2+ are considered indeterminate. All other cases (0 or 1+) were considered HER2-negative (Figure 1(f)). For Ki-67, the IHC score was determined using the *St. Gallen International Expert Consensus*.[12] Tumor cells were evaluated for Ki-67 and scored with the percentage of positively stained nuclei. A cut-off point >20% was considered high ("positive") for Ki-67, while values ⩽20% were considered low ("negative") (Figure 1(g)).

The TMA slides were read by a breast specialist pathologist (MSG) who read the first core of each case. If it was impossible to read the first core due to selection errors or loss of material during the procedure, the second core was analyzed. Informative cores were those that allowed the pathologist to interpret and determine the IHC score successfully in TMA. When cores were missing or in the absence of a tumor, they were considered non-informative. A second breast specialist pathologist (DMU) evaluated the TMA slides independently to assess agreement between observers. To assess intratumoral heterogeneity, two cores from the same case introduced into the TMA were evaluated in a randomly selected subset of cases ($n=12$). The combined analysis of the four biomarker readings on the TMA described before was used to determine the BC subtype in each case. The IHC scores for ER, PR, HER2, Ki-67, and the BC subtype resulting from the TMA reading were compared to those obtained in the original pathology report for the respective case by consulting the medical records. In cases of disagreement, the original slide of the case was re-analyzed by the leading pathologist (MSG) to determine the final IHC score.

Statistical analysis

The sample size was calculated using data from Hospital de Clínicas in Porto Alegre, considering a proportion (P) of positivity of 60% for PR/RE, 20% for HER-2-enriched, and 20% for triple-negative BC. The estimation precision (D) used considered the spectrum of the 10% confidence interval, with semi-amplitude (0.05 above or 0.05 below) as the maximum acceptable error. The confidence interval used was 95% ($Z=1.96$, for $\alpha=0.05$). By applying the formula $N=Z^*Z\ (P\ (1-P))/(D^*D)$, the N of 96 samples were obtained. All statistical analysis was carried out using SPSS version 18 (SPSS IBM, New York, NY, USA). The agreement between the IHC score in the TMA versus the medical records and between the different observers was determined by calculating Cohen's kappa. Sensitivity, specificity, disease prevalence, positive and negative predictive value, and accuracy are expressed as percentages and in Clopper-Pearson confidence intervals.[13] p-Values of and less than 0.05 were considered statistically significant. We consistently followed the STARD2015 as the appropriate reporting guidelines when preparing our manuscript and submitted the completed checklist as Supplemental Material.[14]

Results

TMA performance

In order to incorporate the 242 duplicate BC cases, we constructed 10 TMA blocks, each containing 2 cores of 2.0 mm per case. Each BC case contributes 4 cores (one for each antibody: ER, PR, HER2, Ki-67), resulting in a total of 968 cores. These 968 cores represent the total number of potential cores to be scored in the 10 TMA slides (242 cases × 4 markers).

Regarding the overall quality of the TMA, the immunostaining on the TMA slides showed consistent results with no discrepancies between central and peripheral nuclei. The proper alignment of the nuclei, the inclusion of positive and negative controls, and the orientation of nuclei (such as the placenta) ensured an effective and safe reading by the pathologists.

In TMA slides IHC evaluation, out of the total 968 cores, 97% (940) provided informative results, showing high immunostaining quality and sufficient tumor cellularity (>100 tumor nuclei per core) for adequate scoring (Table 1, a). In 91% of cases, the reading of the first core of the duplicate was sufficient to determine the IHC score. However, in 79 cases, the second core had to be assessed to complete the analysis, highlighting the importance of including duplicate tumors in the TMAs. Uninformative cores were minimal at 2.9% (Table 1, b), primarily due to errors in tumor area selection where both cores lacked tumor tissue. Loss of both cores during processing occurred in only 1% of cases. Importantly, there were no differences observed in the quality of TMA slides stained with different antibodies.

Inter-examiner variability and intratumoral heterogeneity

For all the antibodies evaluated, there was almost perfect and statistically significant agreement in determining the

Table 1. TMA performance and analysis of inter-examiner variability.

Criteria	Total	ER	PR	HER2	Ki-67	*p*-value
a. Informative cores	940/968	235/242	236/242	234/242	235/242	NS
First core	861	218	219	214	210	—
Second core	79	17	17	20	25	—
b. Non-informative cores	28/968	7/242	6/242	8/242	7/242	NS
Loss of both cores	8	2	1	3	2	—
Tumor absent in both cores	20	5	5	5	5	—
c. Inter-examiner variability	908/940	229/235	227/236	223/234	219/235	—
Percentage agreement		97	96	95	93	—
Kappa (95% CI)		0.91 (0.86–0.96)	0.86 (0.77–0.96)	0.85 (0.78–0.93)	0.86 (0.79–0.94)	—
p-Value		<0.001	<0.001	<0.001	<0.001	—

ER: estrogen receptor; PR: progesterone receptor; NS: statistically non-significant difference; TMA: tissue microarray; HER2: human epidermal growth factor 2; 95% CI: 95% confidence interval.

Table 2. Comparison between the results of the IHC score obtained by reading the TMA versus the original report.

Antibody	*n*	Agreement (%)	Disagreement		Kappa (95% CI)	*p*-Value
			False positive[a] (%)	False negative[b] (%)		
a. All	940	871 (93)	25 (2.6)	37 (3.9)	0.81 (0.59–0.98)	<0.001
b. ER	235	226 (96)	5 (2.1)	4 (1.9)	0.88 (0.80–0.97)	<0.001
PR	236	224 (95)	6 (2.5)	6 (2.5)	0.86 (0.79–0.90)	<0.001
HER2[c]	234	210 (90)	6 (2.5)	11 (4.5)	0.71 (0.60–0.82)	<0.001
Ki-67	235	211 (90)	8 (3.3)	16 (6.7)	0.79 (0.70–0.90)	<0.001

IHC: immunohistochemistry; ER: estrogen receptor; PR: progesterone receptor; TMA: tissue microarray; HER2: human epidermal growth factor 2; 95% CI: 95% confidence interval.
[a]Negative in the original report and positive in the TMA.
[b]Positive in the original report and negative in the TMA.
[c]Seven cases (3% among HER2 cases) classified as "false-indeterminate" in the TMA were counted as discordant in the statistical analysis.

IHC score by reading the TMAs by two different pathologists. Kappa values ranged from 0.85 for HER2 to 0.91 for ER (Table 1, c), classified as "almost perfect" by Cohen's criteria.

With regard to intratumoral heterogeneity, the agreement between the IHC scores assigned to the two cores from the same case included in the TMA varied by antibody. For ER and HER2, agreement was almost perfect (100%), with kappa values of 1.0 for both markers. For PR and Ki-67, there was less agreement, classified as moderate for PR ($k = 0.47$) and substantial for Ki-67 ($k = 0.68$). Among the discordant cases, two PR-positive cases in the first core were assessed as negative in the second, and two Ki-67-high cases in the first core were classified as low in the second.

Comparison of TMA results versus original report

Overall, there was a high agreement between the IHC scores obtained in TMA cores and those in the original report, based on the evaluation of the whole section. In the first analysis, 828 of the 940 (88%) IHC scores were concordant, and 112 were discordant. The discordant cases had the original slide containing the whole section reviewed by the study's leading pathologist (MSG), who then reissued the final IHC score. Forty-three IHC scores with initially discordant results were considered concordant after a whole section review using the same immunostaining interpretation criteria. Thus, final agreement was observed between the TMA *versus* the original report in 871 of the 940 IHC scores (93%) evaluated, being classified as almost perfect and statistically significant ($k = 0.81$, $p < 0.001$) (Table 2, a).

Some differences could be observed when the concordance rates were compared among the antibodies (Table 2, b). There was an almost perfect agreement for ER and PR, while for HER2 and Ki-67, this was slightly lower and classified as substantial. The final comparative analysis of the 69 discordant IHC scores showed that in the evaluation of ER and PR, there was a lower and similar frequency of false-positive and false-negative cases in the TMA. For HER2 and Ki-67, there were more discordant cases, with a higher frequency of false negatives (4.5 and 6.7%, respectively) than false positives in the TMA. The most significant discrepancy in results was observed for Ki-67, where

Table 3. Comparative analysis of immunohistochemical evaluation by the traditional method versus TMA (24 cases): time and cost per case.

Item	Traditional procedure		TMA (24 cases)	
	Time (h)	Cost (USD)	Time (h)	Cost (USD)
IHC slide				
ER	8[a]	15.66	8[a]	15.66
PR		11.04		11.04
HER2		11.26		11.26
Ki-67		11.44		11.44
TMA assembly[b]	0	0	3	26.51
Pathologist[c]	0.25	4.21	2	34.14
Total (per case)	8.25	53.61	0.5h	4.58

Exchange rate, 1 USD = 4.98 BRL (Brazilian reais). IHC: immunohistochemistry; ER: estrogen receptor; PR: progesterone receptor; TMA: tissue microarray; HER2: human epidermal growth factor 2.

[a]The ER, PR, HER2, and Ki-67 slides were processed simultaneously. Cost per slide includes the antibody, materials, and labor.

[b]Cost per TMA includes 3 h of labor (trained technician) and materials.

[c]In the traditional procedure, we considered 15 min for evaluating the four markers. In the TMA, we consider 30 min for evaluating the TMA section with 24 cases per marker (30 min × 4 antibodies = 2 h). Value of the pathologist's working hour = 17.00 USD.

24 of the 235 IHC scores were discordant in the TMA compared to the original report.

Working time and cost analysis

Table 3 shows a comparative analysis of the time and cost spent on the technical procedures and evaluation of results using the TMA versus the traditional procedure. The TMA approach reduced the time of IHC evaluation (for the four markers) from 8.5 to 0.5 h per case. This estimated time included glass slide preparation, TMA assembly, IHC staining, and the pathologist's IHC scoring process of an individual or TMA glass slides. Considering the current values, the cost of the IHC panel with four biomarkers (including labor and materials) is $53.61 per case compared to $4.58 spent per case in the TMA approach, a reduction of approximately 11 times. In a TMA of 24 cases, the apparent saving is $1146.52 in total or $47.77 per case.

BC subtyping

Defining the BC subtype is of great clinical relevance in therapeutic management and disease outcomes. Overall, BC subtyping was possible in 97% (237/242) of the cases using the traditional method compared with 89% (217/242) using TMA. In 20 cases (8.4%), the IHC subtype could not be determined due to failure to read the IHC score on the TMA for one or more of the biomarkers analyzed. Between the 217 remaining cases, there was agreement in the BC subtype in 162 (75%) by the two methods. Table 4 presents a detailed analysis of the sensibility, specificity, and overall accuracy of TMA in BC subtyping. Among the 55 discordant cases, 41 (74%) were luminal tumors classified incorrectly as luminal A or B, 5 were HER2 tumor classified incorrectly as luminal A, and 4 triple-negative tumors were incorrectly classified as luminal A, luminal B, or HER-2 subtypes using TMA.

Discussion

In the present study, we propose using a TMA constructed with two 2.0 mm diameter cores of tumor tissue as an alternative to the traditional procedure for the IHC evaluation of ER, PR, HER2, and Ki-67 in BC. The study results suggest that TMA is a fast, highly accurate, and cost-effective method for testing individual BC biomarkers. However, based on the combined analysis of the four antibodies, we do not recommend using TMA with two cores for BC subtyping unless a reduction in costs is necessary to continue testing patients and providing them with treatment.

Regarding TMA feasibility and overall performance, we observed that after an initial and time-consuming period of training for the technical staff, high-quality TMAs were constructed in our laboratory in a satisfactory way. There was a high retention rate of informative cores in the TMAs, with only 2.9% lost (Table 1, a and b), similar to that reported in a previous study.[15] The absence of a tumor is the main cause of non-informative cores, probably due to an error in selecting the area to be punctured in the original block, with the capture of more peripheral cores where the tumor may not be represented. This finding reflects the need for attention and adequate training for the tumor selection stage. The loss of the two cores during the IHC process occurred in only 1% of cases, a frequency similar to that observed in previous studies[16,17] and lower than the 10% loss reported by Visser et al.,[18] who used 0.6 mm cores. Thus, our practice of using 2.0 mm cores in

Table 4. Accuracy of TMA in breast cancer subtyping.

Subtype	*n* (%)	Agreement	False-pos	False-neg	Sensitivity (95% CI)	Specificity (95% CI)	PPV (95% CI)	NPV (95% CI)	Accuracy (95% CI)
All cases	217 (100)	162	19	36	—	—	—	—	—
Luminal A	91 (42)	76	8	7	91.6 (83.4–96.5)	94.0 (88.6–97.4)	91.7 (85.0–95.6)	93.9 (88.3–96.9)	93.0 (88.7–96.0)
Luminal B	85 (39)	59	6	20	74.7 (63.6–83.8)	95.6 (90.8–98.4)	91.6 (83.2–96.0)	85.3 (80.1–89.6)	87.4 (82.3–91.6)
HER2	13 (6)	7	1	5	58.3 (27.7–84.8)	99.5 (97.3–100)	88.4 (50.5–98.3)	97.4 (95.0–98.6)	97.0 (93.8–98.9)
Triple-negative	28 (13)	20	4	4	83.3 (62.6–95.3)	97.9 (94.8–99.4)	85.7 (69.1–94.1)	97.5 (94.1–99.0)	96.0 (92.5–98.2)

pos: positive; neg: negative; PPV: positive predictive value; NPV: negative predictive value; 95% CI: 95% confidence interval; TMA: tissue microarray; ER: estrogen receptor; HER2: human epidermal growth factor 2.

duplicate seems to be ideal to avoid the need to recolor the entire slide due to the infeasibility of analyzing non-informative IHC scores in the TMA.

In microscopic TMA analyses, when the 4 antibodies were analyzed individually (core-by-core), the comparison of 940 IHC scores in the 242 cases showed a high overall agreement (93%, almost perfect) between the TMA results and the original report. In general, an accuracy rate of 90% or above is often deemed high and acceptable for clinical implementation. Many widely accepted diagnostic tests, such as mammography for BC screening, often have sensitivity and specificity rates in the range of 80%–90%.[19] In summary, a 93% accuracy rate is considered high and reliable for clinical practice, meeting or exceeding the standards of well-established diagnostic methods.

However, as described by other authors,[20,21] our study confirms that the TMA performance is not the same for all antibodies. The concordance of IHC scores was higher for ER and PR and lower for HER2 and Ki-67. Our hypothesis to explain these differences is mainly based on intratumoral heterogeneity, which has also been reported as a limitation of the use of TMA in routine IHC evaluation.[22–25]

To investigate this hypothesis, we performed an intratumoral heterogeneity analysis and detected that there is a high agreement between the IHC scores obtained by comparing the two cores of the same case for the ER and HER2 markers and slightly lower for PR and Ki-67. So, we suggest that intratumoral heterogeneity partially explains the occurrence of false-negative and false-positive IHC scores observed in our study (4%, 5%, 10%, and 10% for ER, PR, HER2, and Ki-67, respectively). These results align with previous studies, where discordant results in 2%, 7%, and 8% of cases for ER, PR, and HER2, respectively[26] and 18% for Ki-67[25] were also associated with intratumoral heterogeneity. It is well known that increasing the number of cores for each case in the TMA to cover a larger area of the tumor may reduce its impact. Taken together, our results indicate that, especially for HER2 and Ki-67, the addition of more than two cores (2.0 mm each) per case in the TMA or the use of whole-slide staining to IHC analyses should be considered to decrease the chance of discrepant results.

It was expected that the TMA approach would drastically reduce the work time and the cost of evaluating BC markers spent in Brazilian women's healthcare. Indeed, this is the first study to detail potential savings related to implementing TMA technology in Brazil, specifically inside the public health system (SUS). Importantly, we showed a reduction of 17-fold in time and 11-fold in cost of an individual BC IHC scoring, considering labor time and direct and indirect costs. Taken our current demand of 540 requests per year, TMA would allow us to save 2,500,000 USD per year in BC diagnosis, a reduction of 91% in the amount originally spent on this test in our hospital. It is important to remember that cost-effectiveness is directly linked to the volume of tests performed in each laboratory. So, the time spent to gather sufficient cases to fulfill the TMA should be considered to avoid delays in the release of results. An applicable alternative would be to use TMAs with a smaller number of cases (12, 24, or 36 cores per TMA), which could be produced weekly. "Urgent" cases would be processed immediately using the traditional method and reported in less than 48 h. Even so, TMA may not be a viable method for laboratories with a low volume of tests.

However, to properly decide on implementing new technology as an alternative method in clinical practice, it was crucial to know the TMA accuracy in predicting the BC subtype through the combined analyses of the four BC markers. Based on our results, two arguments can demonstrate that IHC analyses of TMA and whole section are not equivalent and their potential for predicting BC subtype: (1) Accuracy and reliability concerns: the results indicate that TMA yielded a lower rate of successful BC subtyping compared to the traditional method. While the traditional method achieved BC subtyping in 97% of cases, TMA only achieved it in 89% of cases. This discrepancy in

success rates suggests that TMA may be less accurate or reliable in determining BC subtypes. Our findings highlight the variability in TMAs accuracy across BC subtypes, with notable differences in sensitivity, specificity, and overall accuracy. Among Luminal A cases, TMA demonstrated the highest agreement rate but showed relatively lower accuracy in identifying Luminal B cases. In the same way, TMA performance was better in identifying triple-negative cases than HER2-positive. These results highlight the limitations of TMA in accurately capturing the heterogeneity of BC in each individual marker, which can potentially be amplified when they are combined to predict the BC subtype, with serious implications for treatment decisions and patient outcomes. (2) Technical challenges and limitations: the inability to determine the BC subtype in 8.4% of cases due to failure to read the IHC score on the TMA indicates technical challenges associated with this method. Despite the high number of informative cores in our TMAs, issues such as inadequate tissue sampling or technical errors during slide preparation, staining, or interpretation significantly impacted the BC subtyping in those cases The entire IHC process may need to be redone, potentially resulting in significant financial and time losses.

When we highlighted luminal tumors, the predominant subtype in the Brazilian population, we observed that they were frequently misclassified using TMA, with 74% of the discordant cases incorrectly classified as luminal A or B. The correct Ki-67 IHC scoring is crucial for distinguishing between luminal A and B tumors, and the misclassification of Ki-67 as "low" ($\leqslant$20%) or "high" ($>$20%) can be the cause of the higher rate of false-positive or false-negative results for the Ki-67 marker and consequently elevated rate of luminal B incorrectly subtyped in TMA (~30% in our study). Our study aligns with the previous showing a high discordance rate of 38% in Ki-67 scoring in TMAs using the same cut-off of $\geqslant$20%. Our data reinforce the existence of possible reproducibility flaws in the Ki-67 evaluation in TMAs depending on the Ki-67 cut-off point applied in the analyses.[27]

Among the HER2 BC group, despite the high specificity, TMA failed to detect 5 in 13 HER2-positive cases, representing the lowest sensibility (58%) compared to the other subtypes. This result differs from the 98% sensitivity detected previously in a similar study[28] that recommends TMA for HER2 subtyping, a rate that we can't confirm in our study. This could be due to differences in the HER2 scoring methods used in both studies and our patient population composition, including all BC subtypes, which need further investigation. Detecting HER2-positive BC accurately is crucial because it significantly impacts treatment decisions and patient outcomes. They tend to be more aggressive than HER2-negative BC, and they require targeted therapy with drugs like trastuzumab or other HER2-targeted therapies. Finally, our data support previous findings that TMA-based IHC results should be used with caution in BC subtype classification, especially when distinguishing luminal A from luminal B and when interpreting findings for HER2-enriched cancers.[29]

Finally, this study has some limitations that should be acknowledged. Technical challenges, such as misclassification of Ki-67 scores and core selection errors, may have contributed to false-negative and false-positive results. Additionally, our findings are based on a single institution's patient population, which restricts their applicability to broader contexts. Future research should involve larger, multi-center cohorts to enhance the reliability of TMA in BC subtyping in clinical practice.

Conclusion

While TMA offers a fast and cost-effective method for testing individual ER, PR, HER2, and Ki-67 biomarkers in BC, caution is needed when using them for BC subtyping. Challenges include tissue loss during construction and varying performance across markers due to tumor heterogeneity. TMAs perform well in identifying certain BC subtypes, like luminal A and triple-negative, but show less reliability in classifying luminal B tumors. Of concern is their lower sensitivity in detecting HER2-positive BC, impacting treatment decisions. Despite the benefits of efficiency and cost, careful consideration of limitations is crucial in clinical practice, requiring further research to optimize TMA use in BC diagnosis and subtyping.

Author's note

Rubia Denise Ruppenthal is also affiliated with the Department of Pathology and Laboratory Medicine at the Schulich School of Medicine & Dentistry, Western University in London, Ontario, Canada.

Declarations

Ethics approval and consent to participate

This study was approved by the Hospital de Clínicas de Porto Alegre Ethics Committee Board, Porto Alegre, Brazil. Brazilian National Health Council-Ethics in Research number CAAE:96100317.40000.5347. The elevated archival time of the samples used met the requirements for waiver of consent as determined by the institutional ethics review board. All samples were used per the institutional ethics committee's approval.

Consent for publication

Not applicable.

Author contribution(s)

Rubia Denise Ruppenthal: Conceptualization; Investigation; Funding acquisition; Writing – original draft; Methodology; Validation; Visualization; Writing – review & editing; Formal analysis; Project administration; Data curation; Supervision; Resources.

Emily Ferreira Salles Pilar: Investigation; Writing – review & editing; Methodology.
Jordan Boeira dos Santos: Investigation; Methodology; Writing – review & editing.
Rafael Correa Coelho: Investigation; Writing – review & editing.
Carina Machado Costamilan Henriques: Investigation; Writing – review & editing; Methodology.
Diego de Mendonça Uchôa: Investigation; Validation; Formal analysis; Data curation; Writing – review & editing.
Marcia Silveira Graudenz: Conceptualization; Investigation; Funding acquisition; Writing – review & editing; Visualization; Validation; Methodology; Formal analysis; Project administration; Data curation; Supervision.

Acknowledgements

The authors acknowledge Gabriela Remonatto for her support during this project. They also thank Paulo Sanches and Paulo Ricardo Oppermann Thomé from the Department of Biomedical Engineering for providing services in TMA construction. The authors are also deeply appreciative of the patients whose samples were used for this study.

Funding

The author(s) disclosed receipt of the following financial support for the research, authorship, and/or publication of this article: This work was supported by the Fundação de Amparo à Pesquisa do Estado do Rio Grande do Sul [grant number PPSUS 03/2017–17/2551-0001.417-0]; And the Fundo de Incentivo à Pesquisa e Eventos (FIPE) of the Hospital de Clínicas de Porto Alegre [grant number 2019-0539].

Competing interests

The author(s) declared no potential conflicts of interest with respect to the research, authorship, and/or publication of this article.

Availability of data and materials

All data generated or analyzed during this study are included in this published article [and its Supplemental Material].

ORCID iDs

Rubia Denise Ruppenthal https://orcid.org/0000-0003-3311-4701
Marcia Silveira Graudenz https://orcid.org/0000-0002-5872-9537

Supplemental material

Supplemental material for this article is available online.

References

1. Ferlay J, Ervik M, Lam F, et al. Global cancer observatory: cancer today. Lyon, France: International Agency for Research on Cancer, https://gco.iarc.who.int/today (2024, accessed March 19, 2024).
2. Ministry of Health. José Alencar Gomes da Silva National Cancer Institute. Estimativa 2023: incidência de câncer no Brasil. INCA, Rio de Janeiro, https://www.inca.gov.br/sites/ufu.sti.inca.local/files//media/document//estimativa-2023.pdf (2023, accessed March 19, 2024).
3. Liedke PER, Finkelstein DM, Szymonifka J, et al. Outcomes of breast cancer in Brazil related to health care coverage: a retrospective cohort study. *Cancer Epidemiol Biomarkers Prev* 2014; 23(1): 126–133.
4. Rosa DD, Bines J, Werutsky G, et al. The impact of sociodemographic factors and health insurance coverage in the diagnosis and clinicopathological characteristics of breast cancer in Brazil: AMAZONA III study (GBECAM 0115). *Breast Cancer Res Treat* 2020; 183(3): 749–757.
5. Simon SD, Bines J, Werutsky G, et al. Characteristics and prognosis of stage I-III breast cancer subtypes in Brazil: the AMAZONA retrospective cohort study. *Breast* 2019; 44: 113–119.
6. Vieira RA da C, Formenton A and Bertolini SR. Breast cancer screening in Brazil: barriers related to the health system. *Rev Assoc Med Bras* 2017; 63(5): 466–474.
7. Kononen J, Bubendorf L, Kallioniemi A, et al. Tissue microarrays for high-throughput molecular profiling of tumor specimens. *Nat Med* 1998; 4(7): 844–847.
8. Moch H, Kononen J, Kallioniemi OP, et al. Tissue microarrays: what will they bring to molecular and anatomic pathology? *Adv Anatomic Pathol* 2001; 8(1): 14–20.
9. Edge SB, Byrd DR, Compton CC, et al. (ed.) *AJCC cancer staging manual*, 7th ed. New York, NY: Springer, 2010.
10. Hammond MEH, Hayes DF, Dowsett M, et al. American Society of Clinical Oncology/College of American Pathologists guideline recommendations for immunohistochemical testing of estrogen and progesterone receptors in breast cancer. *J Clin Oncol* 2010; 28(16): 2784–2795.
11. Wolff AC, Hammond MEH, Allison KH, 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(20): 2105–2122.
12. Goldhirsch A, Winer EP, Coates AS, et al. Personalizing the treatment of women with early breast cancer: highlights of the St Gallen International Expert Consensus on the Primary Therapy of Early Breast Cancer 2013. *Ann Oncol* 2013; 24(9): 2206–2223.
13. MedCalc Software Ltd. Diagnostic test evaluation calculator. Version 22.021, https://www.medcalc.org/calc/diagnostic_test.php (1993, accessed March 13, 2024).
14. Bossuyt PM, Reitsma JB, Bruns DE, et al. STARD 2015: an updated list of essential items for reporting diagnostic accuracy studies. *BMJ* 2015; 351: h5527.
15. Thomson TA, Zhou C, Ceballos K, et al. Tissue microarray for routine clinical breast biomarker analysis: the British Columbia Cancer Agency 2008 experience. *Am J Clin Pathol* 2010; 133(6): 909–914.
16. Thomson TA, Zhou C, Chu C, et al. Tissue microarray for routine analysis of breast biomarkers in the clinical laboratory. *Am J Clin Pathol* 2009; 132(6): 899–905.
17. Sapino A, Marchiò C, Senetta R, et al. Routine assessment of prognostic factors in breast cancer using a multicore tissue microarray procedure. *Virchows Arch* 2006; 449(3): 288–296.
18. Visser NCM, van der Wurff AAM, Pijnenborg JMA, et al. Tissue microarray is suitable for scientific biomarker

studies in endometrial cancer. *Virchows Arch* 2018; 472(3): 407–413.
19. Oeffinger KC, Fontham ETH, Etzioni R, et al. Breast cancer screening for women at average risk: 2015 guideline update from the American Cancer Society. *JAMA* 2015; 314(15): 1599–1614.
20. van Zwieten A. Tissue microarray technology and findings for diagnostic immunohistochemistry. *Pathology* 2013; 45(1): 71–79.
21. Rossing HH, Talman ML, Lænkholm AV, et al. Implementation of TMA and digitalization in routine diagnostics of breast pathology. *APMIS* 2012; 120(4): 341–347.
22. Dekker TJ, ter Borg S, Hooijer GK, et al. Determining sensitivity and specificity of HER2 testing in breast cancer using a tissue microarray approach. *Breast Cancer Res* 2012; 14(3): R93.
23. Dekker TJ, ter Borg S, Hooijer GK, et al. Quality assessment of estrogen receptor and progesterone receptor testing in breast cancer using a tissue microarray-based approach. *Breast Cancer Res Treat* 2015; 152(2): 247–252.
24. Chavan SS, Ravindra S and Prasad MB. Breast biomarkers: comparison on whole section and tissue microarray section. *J Clin Diagn Res* 2017; 11(3): EC40–EC44.
25. Khoury T, Zirpoli G, Cohen SM, et al. Ki-67 expression in breast cancer tissue microarrays: assessing tumor heterogeneity, concordance with full section, and scoring methods. *Am J Clin Pathol* 2017; 148(2): 108–118.
26. Allott EH, Geradts J, Sun X, et al. Intratumoral heterogeneity as a source of discordance in breast cancer biomarker classification. *Breast Cancer Res* 2016; 18(1): 68.
27. Muftah AA, Aleskandarany MA, Al-kaabi MM, et al. Ki-67 expression in invasive breast cancer: the use of tissue microarrays compared with whole tissue sections. *Breast Cancer Res Treat* 2017; 164(2): 341–348.
28. Furrer D, Jacob S, Caron C, et al. Tissue microarray is a reliable tool for the evaluation of HER2 amplification in breast cancer. *Anticancer Res* 2016; 36(9): 4661–4666.
29. Allott EH, Cohen SM, Geradts J, et al. Performance of three-biomarker immunohistochemistry for intrinsic breast cancer subtyping in the AMBER consortium. *Cancer Epidemiol Biomarkers Prev* 2016; 25(3): 470–478.