Pathology – Research and Practice 212 (2016) 911–918

Contents lists available at ScienceDirect

Pathology – Research and Practice

journal homepage: www.elsevier.com/locate/prp

Original article

Her2 immunohistochemical evaluation by traditional microscopy and by digital analysis, and the consequences for FISH testing

Thomas Marcuzzo[a], Fabiola Giudici[a,*], Elisa Ober[b], Clara Rizzardi[a,b], Cristina Bottin[a], Fabrizio Zanconati[a,b]

[a] Department of Medical, Surgical and Health Science – University of Trieste, Strada di Fiume 447, 34149 Trieste, Italy
[b] UCO di Anatomia e Istologia Patologica – Azienda Sanitaria Universitaria Integrata di Trieste, Strada di Fiume 447, 34149 Trieste, Italy

ARTICLE INFO

Article history:
Received 30 June 2015

Keywords:
Breast cancer
Her2
Digital analysis
Immunohistochemistry
Fluorescence

ABSTRACT

Aim: Her2 protein is the key marker determining the choice of Herceptin therapy after a diagnosis of breast cancer. Its evaluation is made in most laboratories by immunohistochemistry, and interpreted by a pathologist using an optical microscope, a process subject to inter-observer variability, particularly for samples scored as equivocal (2+). Software analysis products have been introduced, seeking to reduce this variability. In this study, we compared the results of both traditional evaluation and a specific software package (VISIA Imaging) to those from fluorescent in situ hybridization (FISH).
Materials and methods: We selected 176 cases of invasive breast cancer sampled during 2012–2014 that were classified as equivocal after evaluation of Her2 immunohistochemistry, and that were also evaluated by FISH. Each tissue slide was scanned with a digital D-Sight Fluo 2.0 microscope and analysed with VISIA Imaging S.r.l. software. The final results were categorised as follows: negative (0–1+), equivocal (2+), or positive (3+). Then each result was compared to that obtained by FISH.
Result: The digital method confirmed 85 samples (48.3%) as equivocal (2+), while 23 (15.1%) were reclassified as negative (1+) and 44 (28.9%) as positive (3+). Of the 176 cases, 24 (13.6%) were not suitable for digital analysis (inadequate). Of 67 reclassified cases (1+ or 3+), 62 were in agreement with FISH results (concordance rate 92.5%). The sensitivity and specificity of the digital method were 100% and 82%, respectively.
Conclusion: The application of this analysis software led to an improvement in the interpretation of cases classified as equivocal, decreasing the need for FISH and increasing diagnostic certainty.

1. Introduction

The expression of the human epidermal growth factor 2 (Her2) membrane protein is a key marker in the prognostic and predictive assessment of breast cancer [1]. Its evaluation is generally made using immunohistochemical tests interpreted by a pathologist, and using in situ hybridisation testing for uncertain cases [2]. However, a pathologist's evaluation of optical microscopy is a subjective parameter, open to inter-observer variability [3]. Many efforts have been made to find a more objective procedure, including digital analysis [4–8]. In this study, a digital analysis software product for Her2 immunohistochemical evaluation was tested. The software uses slide images previously acquired with a digital microscope and semiautomatic analysis of the virtual slide, classifying the sample as positive (3+), negative (0–1+), or equivocal (2+). We evaluated whether the application of this software analysis could lead to an improvement in the interpretation of equivocal cases, decreasing the use of FISH. We examined the concordance rate between the results obtained by digital analysis and by FISH.

2. Materials and methods

We selected 176 cases of invasive breast cancer diagnosed during 2012–2014, classified as equivocal after Her2 immunohistochemical slide evaluation. A confirming FISH test was made only on equivocal cases [1]. All slides were selected from the archive of the Institute of Pathological Anatomy and Histology of the Academic-Hospital public-service corporation "Ospedali Riuniti" of Trieste, Italy. Immunohistochemistry was performed after previous tissue fixation (10% neutral buffered formalin) and paraffin embedding.

* Corresponding author.
E-mail address: fgiudici@units.it (F. Giudici).

http://dx.doi.org/10.1016/j.prp.2016.07.016

Table 1
Immunohistochemical procedure.

√ **Antigen retrieval**: application of a heat procedure with PT-Link instrument (Bioptica) and W-CAP EDTA Buffer pH 8.0 (Bioptica):

○ Preheat: heating of the buffer, its temperature increases from room temperature to 65 °C in 20 min (min).
○ Retrieval: the temperature increases to 95 °C and becomes stable for 20 min;
○ Cool: cooling of the buffer from 95 °C to 65 °C in 20 min;
√ **Endogenous peroxidase inhibition**: H_2O_2 at 3% (Bioptica) for 10 min;
√ **Primary antibody incubation**: clone MIB-1 (Bioptica), 1:200 dilution for 20 min;
√ **Secondary system**: Ultravision Quanto large volume detection kit-HRP (ThermoScientific):

○ Enhancer for 10 min;
○ Polymer for 10 min;
√ **Chromogen**: 3,3′diaminobenzidine (DAB Quanto) for 5 min (ThermoScientific);
√ **Counterstaining**: Mayer haematoxylin for 3 min;
√ **Dehydration**: arising sequence of alcohol;
√ **Clarification with xylene**;
√ **Assembling**.

All immunohistochemical tests were performed using an automatic stainer (Lab Vision, Thermo Scientific), following the procedure shown in Table 1. Each step was followed by washing with Tris-buffered saline containing Tween 20 (Thermo Scientific). After July 2014, another washing step with running water was added after counterstaining, to improve the haematoxylin results. Every slide had been evaluated by a pathologist on an optical microscope with a low magnification and classified as equivocal, in agreement with the most recent guidelines of the American Society of Clinical Oncology and the College of American Pathologists (ASCO-CAP 2013). Cases evaluated before the 2013 guidelines cut-off date, which had been classified using the previous set of guidelines [9], were revaluated, and all were confirmed as equivocal. Each slide underwent digital acquisition, using a D-Sight Fluo 2.0 digital microscope (A. Menarini Diagnostics, Firenze, Italy) and was analysed with VISIA Imaging s.r.l. software (version 2.5.0.1, San Giovanni Valdarno, Italy). The digitizing of a slide requires acquisition with a low-magnification lens (4×) to create a slide preview, on which the area of interest to be scanned with a high-magnification lens (20×) is selected. With this procedure, a digital equivalent of the slide is obtained, and it can be displayed at any magnification from 0 to 20×, maintaining an excellent image quality. The importance of using a high-magnification lens is that the analysis software requires a high-resolution image (not less than 20× magnification). Every equivocal case underwent a FISH confirming test (PathVysion Her2 DNA Probe Kit II, Abbot Molecular) evaluated using a fluorescent microscope, following international criteria (ASCO-CAP 2013). Results collected before the adoption of these guidelines were reevaluated, and no changes were seen in the final diagnosis for any case. After July 2014, fluorescent slides were scanned with a D-Sight Fluo 2.0 digital microscope, and analysed with specific software.

Digital slides were divided into categories based on:

- **Type of specimen**: Surgical sample or biopsy.

The difference in the quantity of tumour sample that can be found in the two different types of specimen is important, as is the subsequent handling of samples during all the steps preceeding the immunohistochemical staining to achieve a perfect reaction quality [10].

- **Suitability for digital analysis**: Inadequate or adequate.

Table 2
Parameters for digital analysis.

Parameter	Range	Function
Only cells with nucleus	ON/OFF	Counts only the membrane associated with a detected nucleus.
Threshold	0–255	Detachment between background and nuclei on the haematoxylin channel.
Minimum area	0–6000	Minimum area under which the object is identified as nucleus.
Crumbling	0–255	Settles the sensibility on the membrane separation.
Threshold + vs −	0–255	Settles the threshold of membrane completeness to classify cells as negative (0) or positive (1+,2+,3+).
Threshold 2+ vs 3+	0–255	Settles the average intensity of complete membrane classification as strong (3+) or moderate (2+).

Some of the slides in our collection did not meet the requirements of the analysis software, mostly due to poor quality of the immunohistochemical signal. These cases were classified as inadequate (technical issues).

- **Profile**: A, B, C, D, E, F, G.

The analysis software uses a sequence of parameters that allow evaluation of staining intensity and of the percentage of the cancer cell membrane that is stained (Table 2). An accurate analysis of the slide requires that these parameters be carefully chosen to avoid misclassification of cells. This is directly affected by the quality of the immunohistochemical reaction, and by the intensity of both the haematoxylin staining of the nuclei and the 3,3′diaminobenzidine (DAB) staining and its distribution. Specific profiles have been developed based on slides with the same staining quality, with reference to a representative slide, from A to G. Every scanned slide was evaluated at low magnification and entered in the appropriate analysis category. A specific profile (F) was developed only for biopsy, even if the staining of the slide is identical to the surgical specimen that underwent B profile analysis. This is due to the different nature of the two samples, biopsies have fewer tissue quantity and different histological tissue treatment (fixation and processing) compared to surgical specimens, causing staining differences that the analysis software can detect, as they are not visible to the naked eye. It is really important to separate the two profiles to avoid misclassification of the slide. Overall, seven different profiles were defined (Table 3), each with its representative slide (Image 1), to make the analysis reproducible and reliable.

- **Staining distribution**: Homogeneous, heterogeneous.

DAB staining intensity distribution is an essential parameter in the choice of the analysis area. Each slide was visualized at low magnification (max 4×) and the staining intensity was evaluated. If the distribution was homogeneous, the analysis area was chosen randomly on the slide. If the distribution was heterogeneous, two separate areas were chosen, one in the most intense region (hot spot), another in the weakest, containing approximately an equivalent number of cells. These areas were analysed separately and a final average result from both was considered.

When these parameters were set, the image analysis was performed using a 20×magnification. It was decided to limit the area of analysis to 20 mm^2 (4× field), due to the time-consuming nature of the analysis (90–120 min). Analysis of the whole section would

Table 3
Profile for the digital analysis.

Profile	Parameters	Value
A	Only cells with nucleus	OFF
Grey nuclei, weak	Threshold	1
intensity	Minimum area	10
Representative slide	Crumbling	25
B13-1404 I	Threshold + vs –	137
	Threshold 2+ vs 3+	77
B	Only cells with nucleus	OFF
Intense blue nuclei	Threshold	40
Representative slide	Minimum area	120
B14-8580/1B	Crumbling	25
	Threshold + vs –	167
	Threshold 2+ vs 3+	87
C	Only cells with nucleus	OFF
Faint slide, soft DAB	Threshold	1
Representative slide	Minimum area	10
B12-772 E	Crumbling	25
	Threshold + vs –	172
	Threshold 2+ vs 3+	102
D	Only cells with nucleus	OFF
Faint slide,	Threshold	1
well-defined DAB	Minimum area	20
Representative slide	Crumbling	10
B12-9154 G	Threshold + vs –	187
	Threshold 2+ vs 3+	117
E	Only cells with nucleus	OFF
Grey/blue nuclei,	Threshold	1
cytoplasmic or	Minimum area	10
intense DAB	Crumbling	25
Representative slide	Threshold + vs –	140
B12-13575/2C	Threshold 2+ vs 3+	65
F	Only cells with nucleus	OFF
Biopsy, Intense blue	Threshold	40
nuclei	Minimum area	10
Representative slide	Crumbling	10
B14-9699	Threshold + vs –	140
	Threshold 2+ vs 3+	60
G	Only cells with nucleus	OFF
Strong intensity mix	Threshold	1
slide (A/E)	Minimum area	10
Representative slide	Crumbling	25
B12-15122/2 o	Threshold + vs –	97
B14-1761	Threshold 2+ vs 3+	35

have taken an excessive amount of time, and the number of cells in an area of 20 mm^2 was expected to be high enough to be representative of the whole section. This was demonstrated to be the case (see Results below). The analysis was based on the identification of DAB-positive cell membranes and on the percentage of the cell membrane stained. Samples were classified into four categories, as follows:

- 0 Negative: no presence of membrane staining, the analysis finds only the haematoxylin staining of the nuclei.
- 1+ Negative: weak and incomplete staining.
- 2+ Equivocal: moderate or incomplete membrane staining.
- 3+ Positive: strong and complete membrane staining.

This classification is equivalent to that recommended in the most recent published guidelines [1]. Every cell is categorized according to these criteria and the software reports the percentage of cells in each category from the total number of analysed cells. The final result is scored based on a positivity cut-off set at 10% of the total cell number, as negative (0–1+), equivocal (2+), or positive (3+). Results from each case were compared to those obtained by earlier FISH tests, which are considered the diagnostic gold standard.

- **Number of cells**: Adequate, inadequate.

When the total number of analysed cells was less than 10,000, the whole analysis was considered unreliable. We observed that with fewer cells the positive/negative percentage was strongly influenced by the number of cells analysed, with small differences leading to large changes. Because this variation started to stabilize above 10,000 cells, we used this number as a reliability threshold. If a single area did not contain enough cells, the analysis was extended until the minimum threshold was reached. If insufficient cells were present within the whole section, the slide was considered inadequate (<10,000). To test the suitability of the selected threshold, we examined the total number of cells analysed for all areas of each slide. All of the analysed areas were included in this examination, thus one for every homogeneous case (129) and two for each heterogeneous case (23). Furthermore, we also examined cell numbers on slides scored as inadequate (<10,000), for a final total of 191 analysed areas.

3. FISH

As previously described, the confirming FISH test needs its analysis with a fluorescent microscope. The analysis started with an observation of the slide by the pathologist with a DAPI filter. This allowed him to evaluate the nuclei morphology and eventual nuclei overlapping, choosing the suitable one. Then the DAPI filter must be switched with the SpectrumOrange to detect Her2/neu signals and count them, losing the nuclear morphological detail that the observer must keep in mind. Then the filter is switched again to the SpectrumGreen, to count Cep17 signals. Both the counts were signed and the observer moved on another nucleus. This process continued until the achievement of the minimal nuclei number, not less than 20 nuclei [1]. After July 2014, the evaluation of the FISH results was made using the digital equivalent of the slide: we used the D-Sight Fluo 2.0 digital microscope to create digital fluorescent slide, and then analysed them with the specific analysis software made by VISIA. This software is an upgrade of the same used in a recent article [11] about standardization of completely automated FISH analysis. The analysis was made on selected areas after examination of all fluorescent slides. The software scanned the selected fields at 60× magnification. A mercury lamp powered the light directed on the sample and a DAPI (4,6-diamidino-2-phenylindol)-stained image was acquired to identify nuclei, using an automated perfect focus system. The software proceeded to scan signals from probes for both Her2/neu (SpectrumOrange) and CEP (SpectrumGreen), using appropriate filters, and collecting data from a greater number of Z-levels compared to the DAPI images. Level number is adjustable, but for these samples it was fixed at 13, separated from each other by 1 μm. The software compressed the scanned images into one that the user can scroll through, removing background from both probes. The pathologist gained the advantages of observing the slide on a screen, and not directly on the fluorescent microscope, with a high image quality including all filters and without need of focusing [12]. The software automatically highlighted the recognizable nuclei and the red and green probes. The pathologist had to check every nucleus and every signal to control the suitability of the analysis. If needed, he could erase, modify or add nuclei or probes signals with a dedicated manual control system. The software continuously updated its final results based on each probe and nucleus accepted by the pathologist. This allowed a fast, effective, and reproducible analysis of the selected nuclei. It was also possible to store high-quality images of each sample, avoiding loss of sample information due to fluorescence decay. Furthermore, digital analysis allowed the observation of a greater number of nuclei compared to direct optical fluorescent microscopy evaluation. The most recent guidelines [1] changed the minimum number of nuclei that a pathologist must count to give a reliable result, decreasing it from 60 to 20. In this study, we tested for effects of

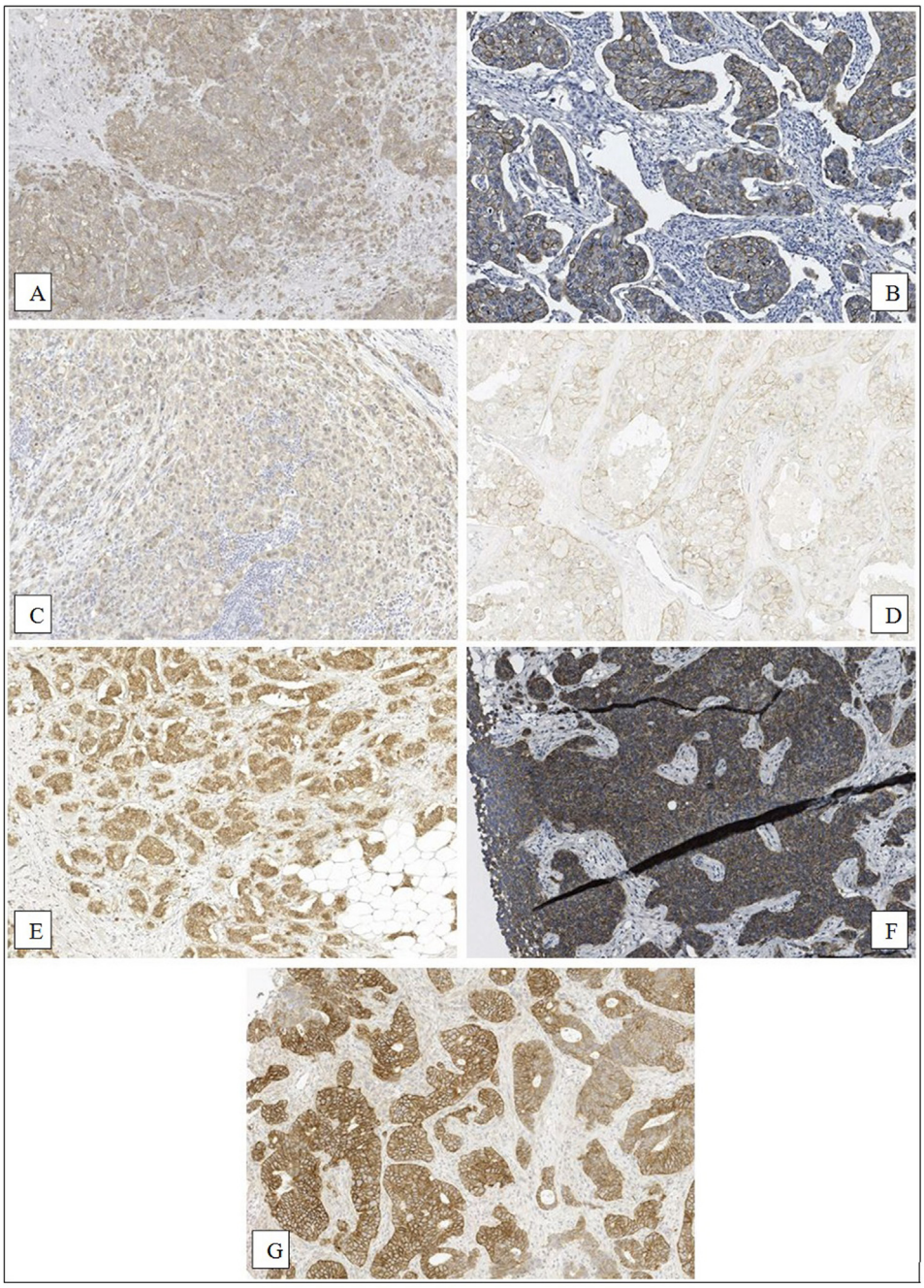

Image 1. Representative slides.

this reduction. Each of the 49 Her2 FISH digitised cases was counted multiple times, starting from a minimum of 20 and increasing the number evaluated until the maximum analysable nuclei within the selected areas was achieved (20, 30, 40, 50, 60, max). The highest number of analysable nuclei ranged from 65 to 384. Results using different cell counts were compared, to determine if it had influenced the final scoring. In these cases, the final result was obtained by agreement of two pathologists, experts in fluorescence interpretation.

4. Results

4.1. Cell number distribution

Cell numbers within the analysed areas ranged from 987 to 127,933, with a mean of 35,483.26; the standard deviation was 25,844.91. Furthermore, 50% of the samples contained between 16,793 and 46,304 cells.

4.2. General results (Table 4)

Of the 176 total cases, 24 were inadequate (12 for technical issues and 12 for cell numbers <10,000), 85 (48.3%) were confirmed as equivocal, 23 and 44 were classified as negative (13.1%) and positive (25.0%), respectively .

In comparison to the gold standard method of FISH the 85 equivocal cases confirmed as equivocal, 67 cases were correctly reclassified as positive or negative, whilst 5 cases were discordant compared to FISH. These cases were classified as amplified by digital analysis and not amplified by FISH. Of all our clinical records (176) the digital analysis of 62 cases is concordant with the FISH results (35%). However, results from five cases (3%) did not agree

Table 4
Characteristics of slides classification after HER 2 immunohistochemistry digital analysis (176 cases).

Characteristic	N	%
Diagnostic Categories		
Inadequate	24	14%
Negative (1+/0)	23	13%
Positive (3+)	44	25%
Equivocal (2+)	85	48%
Total	176	100%
Type of Specimen		
Surgical Sample	132	75%
Core Biopsy	44	25%
Total	176	100%
Suitability for IHC digital analysis		
Inadequate	24	14%
Adequate	152	86%
Total	176	100%
Staining Distribution[a]		
Homogeneous	139	85%
Heterogeneous	25	15%
Total	164	100%
Concordance Rate between IHC Digital Analysis and FISH test[b]		
FISH concordant	62	41%
FISH discordant	5	3%
Equivocal (2+)	85	56%
Total	152	100%

[a] Excluded 12 cases for the impossibility to evaluate the distribution.
[b] Excluded 24 inadequate cases.

with FISH findings These five cases will be described in detail in the discussion.

The 24 samples (14% of cases) that were inadequate for digital analysis were excluded from some of the following investigations, decreasing the clinical records to 152.

4.3. Sensitivity and specificity of the digital method (Table 5)

In the comparison between FISH analysis and digital analysis we excluded the 85 equivocal (2+) cases and also the 24 inadequate ones. Therefore the analysis is focused on 67 cases, of which 39 showed amplification by FISH, while 28 were not amplified. With the digital IHC evaluation, in most cases there was concordance with the FISH evaluation (62 cases of 67), except for five cases (false positives) for which the instrument identified hyperexpression where it was not amplified (positive predictive value of the digital method VPP = 89% [95% CI 78–95%]). The sensitivity and specificity of the digital method were 100% (95% CI [93–100%]) and 82% (95% CI [71–90%]), respectively .

Table 5
Comparison between FISH Analysis and Digital Analysis (excluded 85 equivocal and 24 inadequate cases).

		Digital IHC Analysis		
		Positive (3+)	Negative (1+)	Total
FISH Analysis	Amplified	39	0	39
	Not Amplified	5	23	28
	Total	44	23	67

4.4. Comparison between types of specimen (Table 6)

The records included 132 (75%) surgical specimens and 44 (25%) biopsies. The were no statistical differences between the proportions of inadequate samples in biopsies and in surgical specimens (18% vs 12%, p = 0.47). However, there were a number of samples that remained equivocal after digital analysis, and more of them were found among surgical specimens than in biopsies. Evaluating the concordance percentage with FISH analysis we found it to be slightly higher among biopsy samples but not statistically significant (50% vs 38%, p = 0.27) .

4.5. Comparison between staining distributions (Table 7)

Among the analysis of 164 samples (12 were excluded for the impossibility to evaluate the distribution), 139 (85%) showed homogeneous staining. Furthermore, in both homogeneous and heterogeneous staining categories, there was a high number of equivocal cases. Focusing on the non-equivocal cases, in homogeneous samples, the percentage of FISH-discordant cases was 2% (1 case in 58), while in heterogeneous cases, which were fewer in number, the discordant case percentage was 17% (p = 0.005). The single homogenous discordant case was a biopsy, while all four heterogeneous discordant cases were surgical specimens .

4.6. Comparison among three years (Table 8)

We divided the final results into the three years of the study, and examined them separately. The number of cases classified as equivocal decreased over time, from 50% of total cases in 2012 to 42% in 2014. The percentage of cases with non-equivocal final results, concordant with FISH, increased from 22% in 2012 to 44% in 2014. However, a small percentage of discordant cases was present during all years. The fraction of cases scored as inadequate for technical reasons decreased from 12% to 5% during the three years examined (Image 2) .

Table 6
Digital Analysis according to Type of Specimens.

Characteristic	Surgical Sample	Core Biopsy	p-value	p-trend
Diagnostic Categories				0,43
Inadequate	16 (12%)	8 (18%)	0,47	
Negative (1+/0)	16 (12%)	7 (16%)	0,69	
Positive (3+)	32 (24%)	12 (27%)	0,84	
Equivocal (2+)	68 (52%)	17 (39%)	0,19	
Total	132	44		
Concordance Rate between Digital Analysis and FISH test[a]	Surgical Sample	Core Biopsy	p-value	p-trend
FISH concordant	44 (38%)	18 (50%)	0,27	0,44
FISH discordant	4 (3%)	1 (3%)	1	
Equivocal (2+)	68 (59%)	17 (47%)	0,31	
Total	116	36		

[a] Excluded inadequate cases.

Table 7
Digital Analysis according to Staining Distributions.

Characteristic	Homogeneous	Heterogeneous	p-value	p-trend
Diagnostic Categories				p = 0.83
Inadequate	10 (8%)	2 (8%)	1	
Negative (1+/0)	21 (16%)	2 (8%)	0,45	
Positive (3+)	37 (29%)	7 (28%)	1	
Equivocal (2+)	71 (55%)	14 (56%)	1	
Total	139	25		
Concordance Rate between Digital Analysis and FISH test*	Homogeneous	Heterogeneous	p-value	p-trend
FISH concordant	57 (44%)	5 (22%)	0,07	<0.001
FISH discordant	1 (1%)	4 (17%)	0.005	
Equivocal (2+)	71 (55%)	14 (61%)	0,77	
Total	129	23		

* Inadequate cases were excluded.

Table 8
Result comparison: 2012–2014.

D-sight results	2012		2013		2014	
Inadequate cell N° < 10.000	6	12%	2	3%	4	6%
Inadequate for technical pitfall	6	12%	3	5%	3	5%
Equivocal	26	50%	33	52%	26	42%
FISH concordant	11	22%	24	38%	27	44%
FISH discordant	2	4%	1	2%	2	3%
Total amount	51	100.0%	63	100.0%	62	100.0%

4.7. FISH

Of the 49 cases analysed by FISH with the digital procedure, 41 showed no variation in results based on the different numbers of total nuclei counted. The remaining eight cases did show differences between results obtained from counting the first 20 nuclei and from one other total count category. Moreover, we observed that only the 40-nuclei counts gave the correct result in all cases.

5. Discussion

We compared the results of traditional Her2 IHC evaluation with a specific software package (VISIA Imaging) for Her2 IHC analysis, using FISH as gold standard. To clearly investigate if the application of digital IHC analysis software could lead to a decrease of FISH test and more rapid, reproducible and reliable IHC results, we discuss each step separately.

5.1. Cell number distribution

A minimum of 10,000 cells appeared to be an ideal cut-off for reliability of this analytic data, confirming earlier anecdotal evidence. Indeed, more that 75% of cases in this study provided higher cell numbers, and 50% of samples contained between 16,000 and 46,000 cells. Therefore, we consider the analysis we have conducted to be reliable. Result from cases with fewer cell number should not be considered.

5.2. General results

Some samples may not be suitable for digital analysis, due to characteristics of the sample, as in the case of low cell number previously described, or to technical limitations. It is important to separate these two reasons, as the first is an independent variable (because we cannot determine the number of cancer cells present in a sample) while the other is determined by the quality of the immunohistochemical assay. We have demonstrated that results in 35% of all equivocal cases, with evaluation by optical microscopy, could be better defined by digital analysis avoiding FISH. In a routine use perspective this system will decrease by a third the fluorescence confirming test. The diagnostic contribution that the digital method brings on this matter is undoubtedly relevant.

5.3. Sensitivity and specificity of the digital method

Digital test appeared to have 100% sensitivity, due to the absence of false negatives. However, the presence of FISH-discordant cases

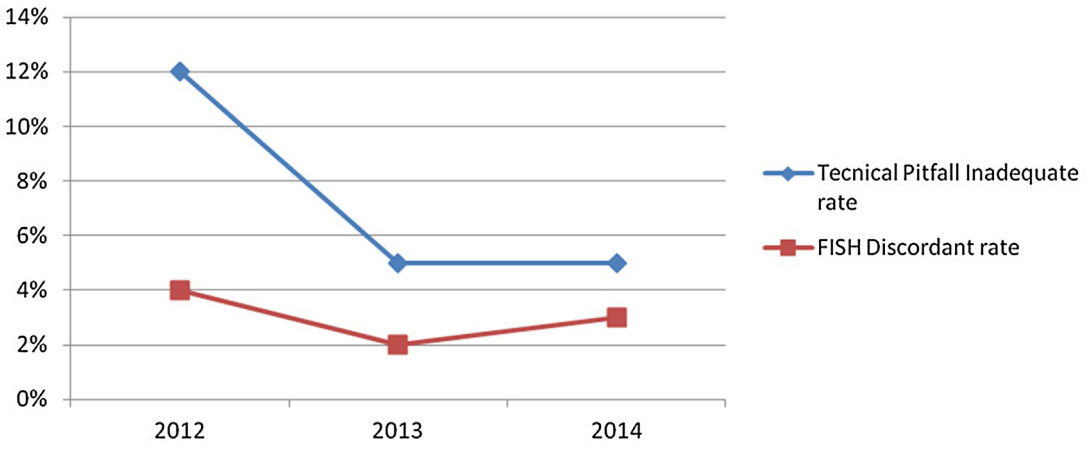

Image 2. Inadequate and Discordant trend during time.

cannot be ignored. In the whole set of samples there were five false-positives which decreased the test specificity to 82%.

5.4. Comparison between types of specimen

The inadequate percentage is similar in both categories, even if it was slightly higher in biopsies. Considering the five FISH discordant cases in detail, four of them were surgical specimens, while one was a biopsy. Biopsies represent a smaller number of samples compared to surgical specimens, but have a slightly higher fraction of inadequate samples. This is inherent to the samples, which present complex challenges to both pre-analytic standardization and immunohistochemical methods. Furthermore, most biopsies samples contain evident artefacts of over- or under-staining due to sample handling, often focused on borders or seen in small-dimension samples. The difference between biopsies and surgical samples immunohistochemistry could be of great influence for some markers and there is variance even for Her2 [13]. This difference has been reported also in other studies on other tissue samples [14]. Guidelines recommend a large group of situations in which IHC test on a single biopsy sample is not adequate and new Her2 test on the surgical specimen must be done [1]. The certainty of correct tissue handling and of IHC method decreases the risks of altered staining and improves digital analysis. A recent study underlined the fact that a percentage of samples scored as positive (3+) by the pathologist, using optical microscopy, were found not to be amplified after FISH test [15]. This shows that there are few samples with high membrane staining that do not have gene amplification. All these factor lead to discrepancies when a biopsy is analysed, even with automated digital software.

5.5. Comparison between staining distribution

Heterogeneous cases present different staining intensities within the same sample.

Frequently it is unclear if variation in intensity is due to focal hyperexpression or is related to artefacts (id: delays in fixation, overfixation, automatic stainer malfunction, or to other reasons related to preanalytical and analytical phases [10,16]). In all heterogeneous cases we selected two areas with differential staining containing comparable cell numbers. Considering the five FISH discordant cases, four of them are surgical specimens that present a heterogeneous staining distribution. Heterogeneous cases should be interpreted with caution.

5.6. Comparison among the years

During the three years of this study, the quality of the immunohistochemical assay changed, due to a strict standardisation of pre-analysis processes and increased control of the immunohistochemical methods. Cold ischemia and formalin fixation time were carefully standardized as suggested by international guidelines [1,9], antibody specificity was periodically tested on positive and negative control tissue, and the analytic activity was submitted to external quality control (NordiQC). Particularly useful was the standardization of haematoxylin counterstaining. A good-quality sample is essential to perform a reliable digital analysis. Lower sample quality risks inaccurate digital results.

5.7. FISH

We observed that limiting the total evaluated nuclei to 20 could lead to doubtful classifications, but also to extend the count to an excessively high nuclei number did not improve the reliability of results. We suggest that 40 is the ideal nuclei number for a reliable count.

6. Conclusion

This analysis software of immunohistochemistry Her2 slide has significant value and utility in the prognostic-predictive evaluation of breast cancer. It gives results needed to patient therapy very rapidly, avoiding the use of FISH on a third of total cases. Digital analysis is also simple to apply, highly sensitive, and reproducible when properly arranged. However, the risks of misclassification, although low, are not entirely removable, and we believe that heterogeneous-staining cases should be confirmed by FISH testing. While the ideal sample for digital analysis is a surgical specimen, if the only available sample is a biopsy, it will be important to be certain of its correct processing and staining. If doubts on its handling are present, FISH may also be required. In the analysis of FISH, a reliably correct evaluation may require increasing the number of nuclei examined to 40. The application of this analysis software can lead to an improved interpretation of equivocal cases, decreasing the need for FISH evaluation, and increasing diagnostic certainty and reproducibility.

Conflict of interest

The authors have no conflict of interest to declare.

Acknowledgements and funding information

This work was supported by grant A. Menarini Diagnostic S.R.L.

The English in this document has been checked by at least two professional editors, both native speakers of English. For a certificate, please see: http://www.textcheck.com/certificate/185gVo. The authors thank Ana Maria Gheorghe-Gutta for her precious contribution in checking the English accuracy of the final version of the manuscript.

References

[1] A.C. Wolff, M.E. Hammond, D.G. Hicks, et al., Recommendations for human epidermal growth factor receptor 2 testing in Breast cancer: American Society of Clinical Oncology/College of America Pathologists Clinical Practice Guideline Update, J. Clin. Oncol. 31 (2013).
[2] C.B. Moelans, R.A. de Weger, E. Van der Wall, P.J. van Diest, Current technologies for Her2 testing in breast cancer, Crit. Rev. Oncol. Hematol. 80 (2011) 380–392.
[3] I. Al Haddabi, A. Qureshi, A. Saparamadu, et al., Inter-observer agreement in reporting HER 2 Neu protein over expression by immunohistochemistry, Indian J. Pathol. Microbiol. 57 (2014) 201–204.
[4] H. Masmoudi, S.M. Hewitt, N. Petrick, K.J. Myers, M.A. Gavrielides, Automated quantitative assessment of HER-2/neu immunoistochemical expression in breast cancer, IEEE Trans. Med. Imaging 28 (6) (2009).
[5] H.R. Ali, M. Irwin, L. Morris, et al., Astronomical algorihms for automated analysis of tissue protein expression in breast cancer, Br. J. Cancer 108 (2013) 602–612.
[6] T. Keay, C.M. Conway, N. O'Flaherty, et al., Reproducibility in the automated quantitative assessment of Her2/neu for breast cancer, J. Pathol. Inform. 4 (2013) 19.
[7] V.J. Tuominen, T.T. Tolonen, J. Isola, ImmunoMembrane: a publicly available web application for digital image analysis of HER2 immunohistochemistry, Histopathology 60 (2012) 758–767.
[8] G. Turashvili, S. Leung, D. Turbin, et al., Inter-observer reproducibility of Her2 immunohistochemical assessment and concordance with fluorescent in situ hybridization (FISH): pathologist assessment compared to quantitative image analysis, BMC Cancer 9 (2009) 165.
[9] A.C. Wolff, M.E. Hammond, D.G. Hicks, et al., American Society of Clinical Oncology/College of American Pathologists Guideline recommendations for human epidermal growth factor receptor 2 testing in breast cancer, Arch. Pathol. Lab. Med. 25 (2007).
[10] G. Bussolati, E. Leonardo, Technical pitfalls potentially affecting diagnoses in immunohistochemistry, J. Clin. Pathol. (2008) 1184–1192.
[11] E.M.J. Van der Logt, D.A.J. Kuperus, J.W. van Setten, et al, Fully automated fluorescent in situ hybridization (FISH) staining and digital analysis of He in breast cancer: a validation study, PLoS One 10 (4) (2015) e0123201, http://dx.doi.org/10.1371/journal.pone.0123201.
[12] https://140.105.126.6:443/webdsightplus/instantviewer/hash/7db8a8c983e5b6b86d2cc5bd42048fe9.

[13] C.S. Connor, O.W. Tawfik, A.J. Joyce, et al., A comparison prognostic tumor markers obtained on image-guided biopsies and final surgical specimens, Am. J. Surg. 184 (October (4)) (2002) 322–324.
[14] L. Taillade, F. Penault-Llorca, T. Boulet, et al., Immunohistochemichal expression of biomarkers: a comparative study between diagnostic bronchial biopsies and surgical specimens of non-small-cell lung cancer, Ann. Oncol. 18 (June (6)) (2007) 1043–1050.
[15] Z. Varga, A. Noske, C. Ramach, B. Padberg, H. Moch, Assessment of HER2 status in breast cancer: overall positivity rate and accuracy by fluorescence in situ hybridization and immunohistochemistry in a single institution over 12 years: a quality control study, BMC Cancer 13 (615) (2013).
[16] A. Babic, I.R. Loftin, S. Stanislaw, et al., The impact of preanalytical processing on staining quality for H&E, dual hapten, dual color in situ hybridization and fluorescent in situ hybridization assays, Methods 52 (2010) 287–300.