PATHOLOGICA 2018;110:83-91

ORIGINAL ARTICLE

A National Quality Assurance program for breast immunohistochemistry: an Italian perspective

E. GUADAGNO[1], G. DE ROSA[1*], O. NAPPI[2]
[1] Department of Advanced Biomedical Sciences, Pathology Section, University of Naples Federico II, Naples, Italy; [2] Pathology Unit, A. Cardarelli Hospital, Naples, Italy; *past President of the Italian Society of Anatomic Pathology and Cytology

Key words

Immunohistochemistry • HER2 • Quality control • Italian

Summary

A national project for the quality assessment of breast immunohistochemistry, involving 155 pathology laboratories distributed all over the Italian territory (19 regions), was carried out. The Project lasted one year from December 2014 to December 2015 and it was strongly supported by the Italian Society of Anatomic Pathology (SIAPEC/IAP). Proficiency tests were carried out by the Nordic Immunohistochemical Quality Control (NordiQC) organization. The main aim of the project was to investigate on the general performance of immunohistochemistry (ER, PR and HER2) in the field of breast cancer in the Italian territory, in order to emphasize any difference and give practical support to laboratories in daily practice.
The present review article focused on the description of this extraordinary pioneer Italian experience. Besides NordiQC results, further analysis concerning epidemiology and geographical distribution were done.
Aim of the study was to analyze the general results and to discuss on the benefits that a national quality control program may have if it became a mandatory service provided by the National Health Care System.
In general, the Italian data were in accordance with the general results obtained from the "official" NordiQC HER2, PR and ER assessments. A HER2 scoring consensus between labs and assessor group was achieved in 80% of cases.
Interestingly, what emerges from our study is that no substantial differences exist among the three Italian macro-areas (North, Center and South) in the quality of Immunohistochemistry performed for breast cancer. No statistically significant difference was even found between laboratories that perform more or less than 100 tests/year.

Introduction

Standardization is a tool through which to ensuresafety, reliability and good quality in laboratory activity and pursuingit is essential in everyday medical practice [1]. In the field of anatomic pathology, immunohistochemistry(IHC) can no longer be considered only a diagnostic tool; it must be taken as a means for medical acts because of its prognostic and predictive properties [2]. Three main fields, each including numerous steps, take part of this technique and they all need to be standardized: pre-analytical, analytical and post-analytical. Recommendations have been issued involving each of them [3-7].
To standardize means to conform to a model. Many external proficiency testing (PT)programs [8-10] aim to issue recommendations for the analytical steps useful to pursuit this scope. About one-third of laboratories that participate in these programs achieve optimal results, while one-third are 'good' and one-third fail [2 11 12], indicating that PT programs still have a long way to go.
In breast cancer, quantitative immunohistochemical evaluation of Estrogen Receptor (ER), Progesterone receptor (PR), HER2 and Ki67 is essential to address the therapeutic strategy [13]: identifying hormone responsive and HER2 amplified tumours may ensure patients a life-saving therapy or spare them foruseless treatments in case of negative results. It follows that reducing the total number of false positive and false negative results is paramount in order to achieve the maximum of cures at the lowest price in terms of costs and side-effects [14].
Italy is the 4th-largest national economy in Europe. In 2000 Italy's Healthcare System was regarded, by World Health Organization's ranking, as the second best in

Acknowledgments
The photos and technical data have been provided by NordiQC.

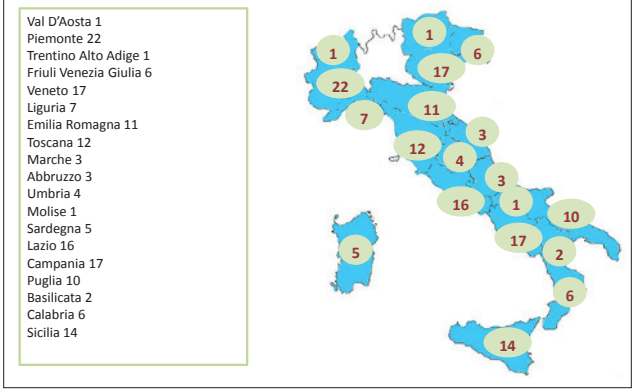
Correspondence
Elia Guadagno, Department of Advanced Biomedical Sciences, Pathology Section, University of Naples Federico II, via Pansini 5, 80131 Naples, Italy - Tel./Fax +39 081 7463433/3475 - E-mail: eliaguadagno84@gmail.com

the world after France [15]. According to the CIA World factbook, Italy has the world's 14th highest life expectancy [16]. The Italian National Outcomes Program [17] permits measurement of variations in the quality and outcomes of care by region: Italy has the largest internal difference of gross domestic product (GDP)per capita between regions of any European country. Although in theory the entire healthcare system operates under one central ministry of health, the national index score of Italy is a mix of Northern Italian and Rome Green scores, and Southern Italian Red scores, resulting in a lot of Yellows (the performance of the respective national healthcare systems was graded on a three-grade scale for each indicator, where the grades had the meaning of Green = good , Yellow = so-so and Red = not-so-good) [17].
In 2015, the Italian Society of Anatomic Pathology and Cytology (SIAPEC)/ International Academy of Pathology (IAP) promoted a National Quality Control Program on breast cancer IHC, involving 158 Italian labs. PTs were carried out by the Nordic Immunohistochemical Quality Control (NordiQC) organization [8 9].
The aim of this study is to present the Italian results of the NordiQCPT, identify regional differences in IHC staining quality and discuss the potential benefits of a National Quality Control program as a mandatory service provided by the National Healthcare System.

Material and methods

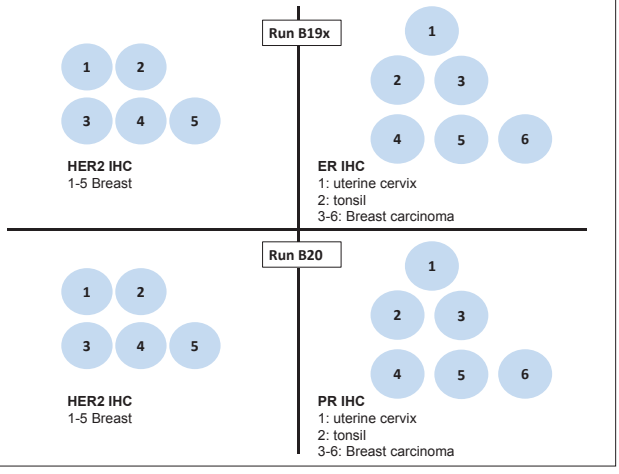
In October 2014, a working group was set up by the SIAPEC/IAP to carry out a National Quality Control program on IHC, enrolling 158 Italian laboratories.19 out of 20 Italian Regions participated (Fig. 1); Lombardiay did not participate because it had already started a similar project independently. Each lab's referent was contacted both by phone and e-mail and was asked to send by e-mail a signed form (inclusive of contacts, names, addresses and number of HER2 tests/year for 2014). SIAPEC/IAP initiated a collaboration with NordiQC to host and perform the IHC PT. An introductory run named run B19X was established for the 158 Italian labs, identical to the "official" NordiQC run B19 (April 2015) assessing ER and HER2. Subsequently, the Italian labs participated in the "official" NordiQC run B20(September 2015) for PR and HER2, together with all other NordiQC laboratories (i.e. 525).
For both run B19X and B20, the participating Italian laboratories completed an on-line questionnaire for the protocols used by the individual laboratory for each of the markers evaluated. The protocol data reported by the laboratories were archived in the NordiQC database and used for the subsequent analysis of the assessment results. Two unstained sections for each marker were circulated to the participating laboratories, in order to perform the IHC assays according to their standard protocols as submitted in the questionnaire.
Slides of tissue microarrays (TMAs) from standard processed formalin-fixed, paraffin-embedded (FFPE) tissue

Fig. 1. Distribution of participating laboratories among Italian regions.



were used. Tissue fixation and processing had been carried out according to the recommendations provided by the American Society of Clinical Oncology/College of American Pathologists (ASCO/CAP) [6 7].
The TMAs for HER2 IHC (run B19X and B20) included 5 cores (4-5 mm) of 5 different breast carcinomas (Fig. 2). HER2 IHC expression levels were definedby NordiQC by using the three FDA approved kits and antibodies, HercepTest™ Dako, Oracle™ Leica and PATHWAY® Ventana. HER2 amplification status was verified by HER2 fluorochrome in situ hybridization (FISH). HER2 IHC expression levels in the TMAs reflected the range seen in a diagnostic setting and thus included a HER2 IHC 3+ tumour with high level gene amplification, a 2+/3+ tumour with low level gene amplification, a 1+/2+ tumour without gene amplification and two 1+/0+ tumours without gene amplification. The TMAs for ER and PR included cores (4-5 mm) of tonsil, uterine cervix and 4 different breast carcinomas (Fig. 2). Tonsil and uterine cervix were primarily used as positive and negative tissue controls, respectively, for the marker evaluated. Breast carcinomas showed diagnostically relevant expression levels ranging from negative to strong.

Fig. 2. Representative scheme of slides with tissue cores.



The slides stained by the participants were assessed by the NordiQC expert panel consisting of consultant breast pathologists and biomedical scientists, all experienced in IHC PT. Each slide was by consensus marked as optimal, good, borderline or poor, primary based on the precision of the IHC result and concordance to the NordiQC established levels of the target analyte and addressing the technical quality of the IHC staining result. Results evaluated as optimal and good were both considered sufficient, while borderline and poor were considered insufficient. For the latter individually tailored suggestions to improve the staining quality were provided by NordiQC. HER2 results were considered optimal when all scores for the 5 breast carcinomas were as expected and good if the 3+ HER2 amplified carcinoma or the negative cases showed a 2+ reaction. No patient was misdiagnosed and additional FISH would identify a correct HER2 status. The result was evaluated as borderline if there was a low signal-to-noise ratio compromising the interpretation and poor in case of false negative results (when a 3+ or a 2+ amplified tumor showed a 1+/0 reaction) or of false positive (when a 1+/0 or a 2+ unamplified tumor showed a 3+ reaction).

For ER and PR, the staining reaction was evaluated as optimal when all scores wereas expected and concordant to the NordiQC reference data. If a reduced proportion of cells was demonstrated in the carcinomas expected to be labelled, but still a positive result was seen in ≥10% of the neoplastic cells, the result was evaluated as good. If a significant reduced analytical sensitivity was observed and ≥ 1 but < 10% of neoplastic cells were identified, the result was evaluated as borderline. In case of a false negative < 1% or false positive staining reaction, the resultwas evaluated as poor.

Data of ER, PR and HER2 results of Italian labs were extrapolated and compared to the overall data of the NordiQC participants available on the web site (www.nordiqc.org).

The analysis of association between the annual number of HER2 tests performed by each lab and the proportion of sufficient results was carried out with Fisher's exact test. The same test was used to verify the presence of any difference in geographical distribution of results by diving the labs into three main macro areas: North (labs from Piemonte, Val D'Aosta, Veneto, Friuli Venezia Giulia, Trentino Alto Adige, Emilia Romagna and Liguria), Center (labs from Toscana, Lazio, Umbria, and Marche) and South (labs from Campania, Sicilia, Puglia, Abbruzzo, Molise, Sardegna and Calabria).

A probability (p) value less than 0.05 was considered statistically significant. All tests were two sided and carried out with GraphPad Prism 5 software (GraphPad Software, La Jolla, CA, USA).

Results

RUN B19X HER2 ASSESSMENT

154 labs returned the slides for the evaluation. Of these 79% achieved a sufficient mark (optimal or good) (Tab. I). In 83% of the labs (128/154) the FDA/CE IVD approved assays were used. PATHWAY based on rabbit monoclonal antibody (rmAb) clone (Ventana) was used by 56 laboratories, gaining sufficient results in 84% of cases; the second most used assay was Hercept Test SK001 (Dako), based on a rabbit polyclonal antibody and it was used by 26 laboratories, with a proportion of sufficient marks of 92% (21/26). In both assays the application of optimal protocol settings (suggested by the vendors) ensured a higher pass rate (92% and 96%, respectively). These results were in accordance with those obtained from the "official" NordiQC HER2 assessment in RunB19. Only in 26 cases laboratory developed (LD) assays were used with a proportion of sufficient marks of 70 % (18/26).

In 29 cases a poor result was observed. It was mainly due to a 1+ reactivity in an amplified 2+/3+ breast cancer core. Such false negative results were observed in both FDA/CE IVD approved (but modified) and LD assays and typically caused by reduced heat-induced epitope retrieval (HIER) time, short incubation time of the primary antibody and/or less sensitive detection kit (Fig. 3). A scoring consensus between labs and assessor group was achieved in 80% of cases: labs with sufficient staining results achieved a scoring consensus in 93 cases out of 113 (82%) while those with insufficient results in 20 cases out of 28 (71%).

RUNB19X ER ASSESSMENT

A total of 151 Italian labs returned the stained slides. In 78 cases (51%) a sufficient mark was achieved (Tab. I). Both concentrated and ready to use (RTU) antibodies were applied. In 101 laboratories, RTU antibodies were used, rmAb clone SP1 (Ventana), in particular (73 labs). Sufficient results were obtained in 95% of cases. Most of poor results (22/29 labs) were obtained with antibody clone 6F11 (Leica/Novocastra) purchased concentrated or in RTU format (6/8 labs) and with antibody clone 1D5 (Dako) concentrated (9/9 labs) or in RTU format (6/8 labs). Demonstration of ER was most challenging in the breast carcinoma core no. 4 (expected weak nuclear staining reaction of 40% of the neoplastic cells) where carefully calibrated protocols were required. Uterine cervical tissue was used as positive control, displaying a moderate to strong nuclear signal in both the squamous epithelium and in the glands but also in stromal cells, except for endothelial and lymphocytic cells. In cases where the positive control stained as expected, core nr. 4 showed reliable reactivity. Core no.3 ER negative breast carcinoma was used to assess the specificity of the protocol: staining of tumour cells in this core was assessed as false positive. The protocol was considered highly sensitive when the signal was detected in normal stromal cells and not only in normal epithelial cells, where a strong signal is expected (Fig. 4).

RUNB20 PR ASSESSMENT

123 Italian labs returned the stained slides. In 108 cases (88%) a sufficient mark was achieved (Tab. II). The best

Tab. I. Assessment marks for Run B19x, including tests with ER and HER2 antibodies.

ER								
Concentrated antibodies	**N**	**Vendor**	**O**	**G**	**B**	**P**	**Suff***	**Suff° OPS**
mAb clone **1D5**	9	Dako	0	0	0	9	0%	-
mAb clone **6F11**	29	Leica/Novocastra	0	3	4	22	10%	-
rmAb clone **EP1**	6	Dako	1	0	0	5	17%	50%
rmAb clone **SP1**	2 1	Thermo/Neomarkers Aczon	0	0	3	0	0%	-
Unknown	2	Unknown	0	0	0	2	0%	-
Ready-to-use antibodies								
mAb clone **1D5 IR/IS657**	8	Dako	0	1	1	6	13%	-
mAb clones **1D5+ER-2-123 SK310**	3	Dako	0	0	0	3	0%	-
mAb clone **6F11 PA0151**	8	Leica/Novocastra	0	1	1	6	13%	-
rmAb **EP1 IR/IS084**	10	Dako	1	2	3	4	30%	60%
rmAb clone **SP1 790-4324/5**	73	Ventana	43	26	4	0	95%	94%
Total	151		45	33	16	57	-	
Proportion			29%	22%	11%	38%	51%	
HER2								
FDA/CE IVD approved HER2 assays	**N**	**Vendor**	**O**	**G**	**B**	**P**	**Suff.**	**Suff. OPS**
PATHWAY rmAb clone **4B5,790-2991**	56	Ventana	45	2	0	9	85%	92%
CONFIRM, rmAb clone **4B5, 800-4493**	10	Ventana	6	2	0	2	80%	89%
CONFIRM, rmAb clone **4B5, 800-2996**	1	Ventana	1	0	0	0	-	-
HercepTest SK001	26	Dako	21	3	0	2	92%	96%
HercepTest K5207	15	Dako	13	1	0	1	93%	93%
HercepTest K5204	7	Dako	6	0	0	1	86%	100%
Oracle mAb clone **CB11, TA9145**	13	Leica	5	0	0	8	38%	50%
Antibodies for laboratory developed HER-2 assays, conc. antibody	**N**	**Vendor**	**O**	**G**	**B**	**P**	**Suff.**	**Suff.OPS**
mAb clone **CB11**	4 3 1	Leica/Novocastra Cell Marque Biocare	4	3	0	1	88%	86%
rmAb clone **SP3**	1	Thermo/NeoMarkers	0	0	0	1	-	-
pAb clone **A0485**	14	Dako	9	2	0	3	79%	83%
Antibodies for laboratory developed HER-2 assays, RTU	**N**	**Vendor**	**O**	**G**	**B**	**P**	**Suff.**	**Suff.OPS**
mAb clone CB11, RTU-CB11	1	Leica/Novocastra	0	0	1	0	-	-
mAb clone CB11, 237M-18	1	Cell Marque	0	0	1	0	-	-
pAb E2441	1	Spring Bioscience	0	0	0	1	-	-
Total	154		110	13	2	29	-	-
Proportion			71%	8%	1%	20%	79%	-

* Proportion of sufficient stains; ° Proportion of sufficient stains with optimal protocol settings only

Fig. 3. Serial sections of the tissue microarray for HER2 used in runs B19x and B20, immunostained in three laboratories. Core A (upper row) is a carcinoma with high HER2 gene amplification, core B (middle row) a carcinoma with a low HER2 gene amplification, and core C (lower row) a carcinoma without HER2 gene amplification. Lab 1 (left column) used an FDA approved kit and obtained optimal results: 3+ staining in core A, 2+ in core B and 1+ in core C. Lab 2 (middle column) used a laboratory developed assay (LDA) with a too weak staining and obtained a 2+ reaction in core A. In a diagnostic setting, this tumour would be reflexed to FISH test for final HER2 status, increasing costs and turnaround time. In coreB the lab obtained a 1+ reaction. In a diagnostic setting this tumour would not be reflexed to FISH, and the patient consequently not offered HER2 targeted therapy in spite of the HER2 gene amplification. This staining result was by NordiQC assesses as poor. Lab 3 (right column) used an LDA with a too strong staining and obtainedin core C a 3+ reaction. In a diagnostic setting this tumour would not be reflexed to FISH, and the patient would be offered an ineffective but costly and potentially hazardous HER2 targeted therapy.This staining result was also assesses as poor.

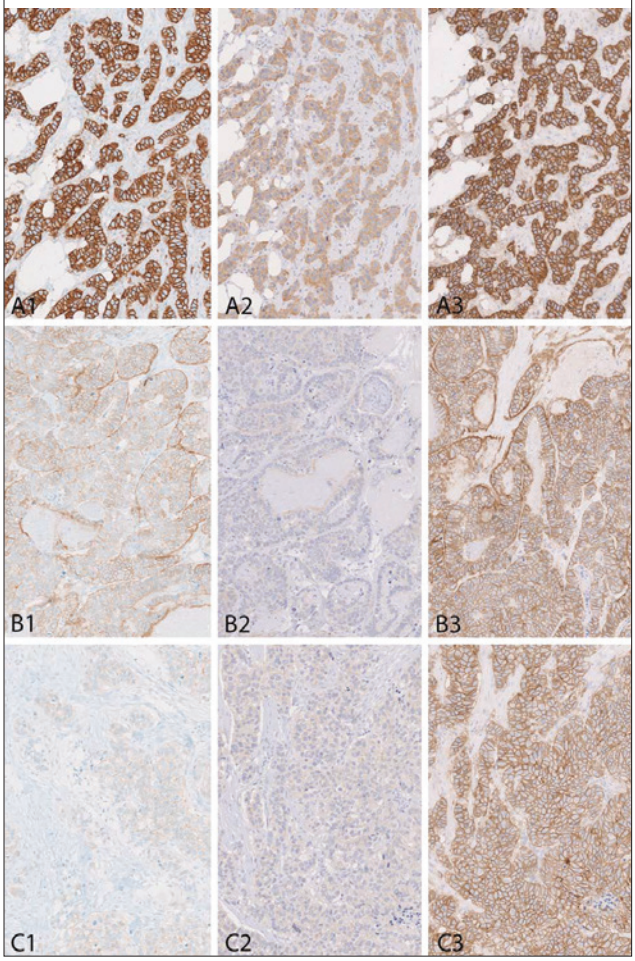
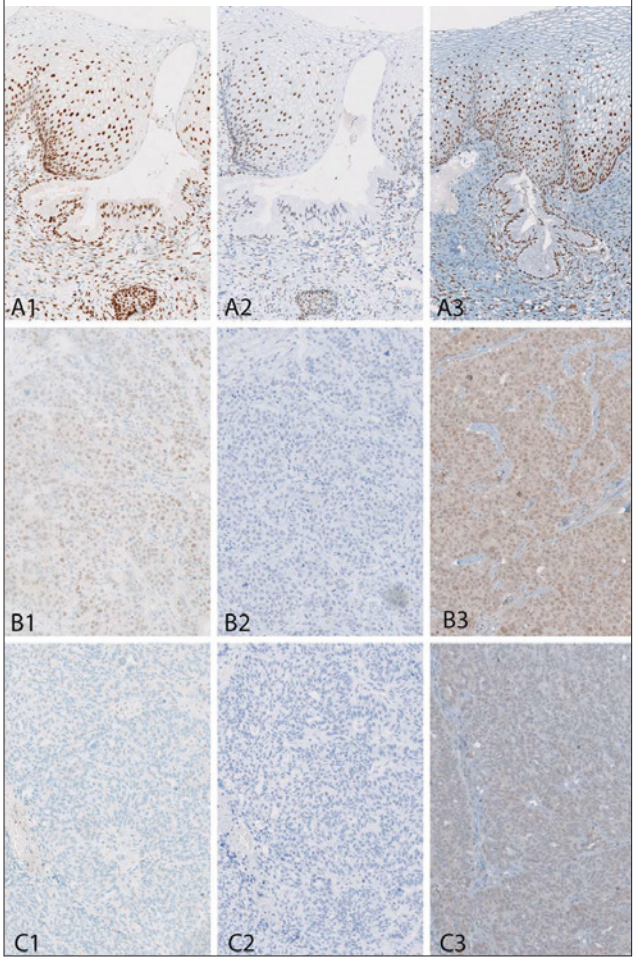


Fig. 4. Serial sections of the tissue microarray for estrogen receptor (ER) used in runB19X, immunostained in three laboratories .Core A (upperr ow) is normal uterinecervix, core B (middle row) is an ER positive breast car cinoma with moderate expression (40-60% positive nuclei), and core C (lower row) is an ER negative breast carcinoma. Lab 1 (left column) obtained an optima staining reaction usinga carefully calibrated protocol based on the rabbit monoclonal antibody clone SP1: Incore A, virtually all squamous and columnar epithelial cells as well as most stromal cells show a moderate to strong nuclear staining reaction, while endothelial cells and lymphocytes are negative. In core B, a heterogeneous positivity of the neoplastic cells is seen, while in core C, the carcinoma i snegative.Lab2 (middle column) also used the clone SP1 but in a too low titre causing a reduced analytical sensitivity: the columnar and basal squamous epithelial cells in core A and the carcinoma in core B are false negative. In a diagnostic setting this patient would not be offered antihormonal therapy. This staining result was therefore assesses as poor. Lab 3 (right column) used the mouse monoclonal antibody clone 1D5 in a too high titre giving a poor signal-to-noise ratio: in core B and especially core C an aberrant cytoplasmic staining reaction hampers the interpretation of the ER reaction. This staining result was therefore assesses as borderline.



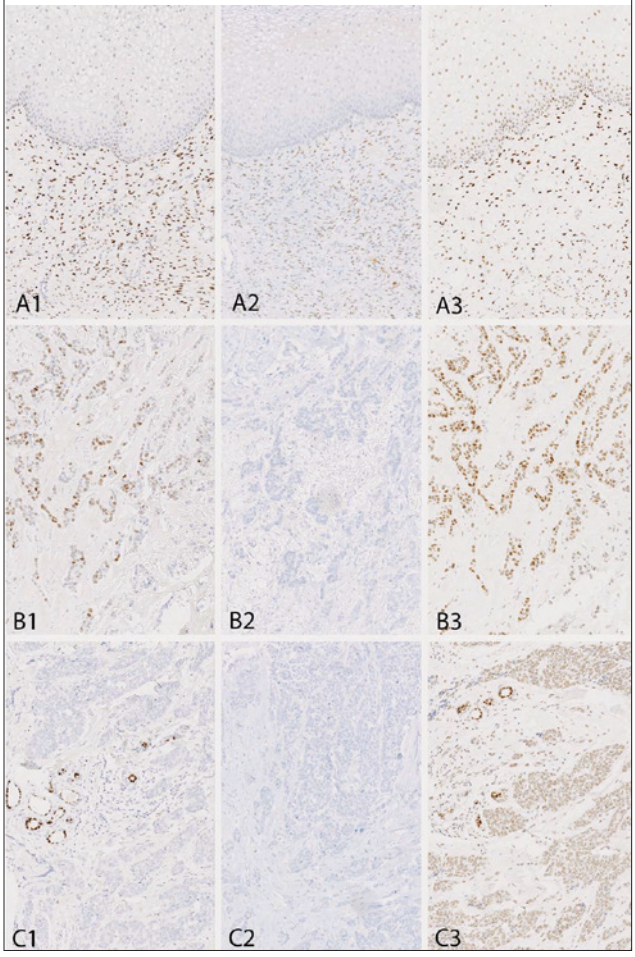
results were obtained with concentrated mouse monoclonal antibodies clone PgR636 (Dako), clone 16 (Leica), their RTU counterparts and clone PgR 1294. In 60 laboratories rabbit monoclonal antibodyclone 1E2 (Ventana) was used, in 48 (80%) a sufficient result was obtained. In most poor cases a false positive nuclear staining was present in ≥ 10% of cells in the tonsil (Fig. 5).

The Italian data were in accordance with the general results of the RunB20 (see www.nordiqc.org).

RunB20 HER2 assessment

A total of 122 Italian labs returned the slides. A pass rate of 79% (96/122) was achieved (Tab. II). Insufficient results were characterized by false negative staining reaction (85%, 22/26), mainly observed as 0/1+ in the 2+ HER2 gene amplified breast carcinoma core. In other insufficient assays poor signal-to-noise ratio, im-

Fig. 5. Serial sections of the tissue microarray for progesterone (PR) used in run B20, immunostained in three laboratories. Core A (upper row) is normal uterine cervix, core B (middle row) is a PR positive ductal breast carcinoma, and core C (lower row) a PR negative ductal breast carcinoma. Lab 1 (left column) obtained an optimal staining reaction using a carefully calibrated protocol based on the rabbit monoclonal antibody clone 16: in core A, the squamous epithelium shows a moderate staining of the basal cells, and all stromal cells (apart from endothelium and scattered lymphocytes) are strongly positive. In core B, a heterogeneous positivity is seen, while in core C, only the normal ducts are positive. Lab 2 (middle column) used a protocol with a too low sensitivity: In core A the basal cells are unstained and in core B the carcinoma is false negative. This staining result was therefore assesses as poor. Lab 3 (right column) used the Ventana RTU product for rabbit monoclonal antibody clone 1E2 but had modified the protocol to make it more sensitive. In core A, staining reaction is now seen in suprabasal epithelial cells and faintly in endothelial cells. In core C, the carcinoma is false positive.This staining result was therefore assesses as poor.



paired morphology or excessive counterstaining made the interpretation difficult. The FDA/CE-IVD approved HER2 assays PATHWAY and CONFIRM from Ventana and HercepTest (SK001) from Dako were the most successful. Laboratory developed assays were less successful.

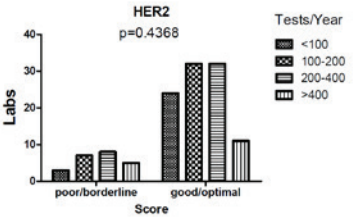
A scoring consensus was achieved in 84% of cases. Italian data were in accordance with the general results of the RunB20 (see www.nordiqc.org).

Fig. 6. Fisher's exact test: no difference in the frequency of sufficient (optimal and good) results for HER2 were observed among the four groups of laboratories (with < 100, or 100-200, or 200-400, or > 400 HER2 tests/year); p = 0.4368 and $\chi^2 = 2.72$.

HER2

Score/tests/year	<100	100-200	200-400	>400
Poor/Borderline	3	7	7	5
Good/Optimal	24	33	32	11

p= 0.4368; **χ2 =2.72**



Comparison of Italian laboratories

Data on the average number of HER2 tests/year were available for 122 laboratories: it ranged from 36 to 1570 tests/year. Laboratories were divided into 4 groups on the base of their performance status (Fig. 6), from < 100 tests/year to > 400 tests/year. In each of them the number of sufficient (optimal/good) and insufficient results was examined. No statistically significant differences were found among these groups (p = 0.4367; χ^2=2.72).

The application of the Fisher's exact test revealed the absence of any statistically significant difference in the proportion of sufficient results of HER2, ER and HER2 score consensus among the three Italian macro areas (North, Center and South) , except for PR that gave more sufficient results in the North than in the South and the Center (Fig. 7).

Discussion

Quantification of immunohistochemical reactivity is an important goal to be pursued in the era of the "next generation IHC", as the implementation of the targeted cancer therapy has introduced a predictive role for tissue markers. Accurate quantification requires method of high quality both in terms of sensitivity and specificity. Several organizations (NordiQC, UK NEQAS ICC, RCPAQAP, cIQc) work internationally and have long time experiences on programs based on expert panel-based qualitative assessment systems. PT in IHC aims to examine the analytical outcome of the staining and to relate any difference to the antibodies, protocol parameters and stainerplat forms 9). The optimization of the analytical phase is a compelling issue also due to intra-observer and inter-observer variability of the interpretation of the staining results. The lack of reproducibility of scoring by pathologists is being addressed by digital algorithms applied to computer assisted image analysis [18-21], whose effects have yet to be explored.

Tab. II. Herein are reported the assessment marks for Run B20 modified for Italian laboratories only, including tests with PR and HER2 antibodies.

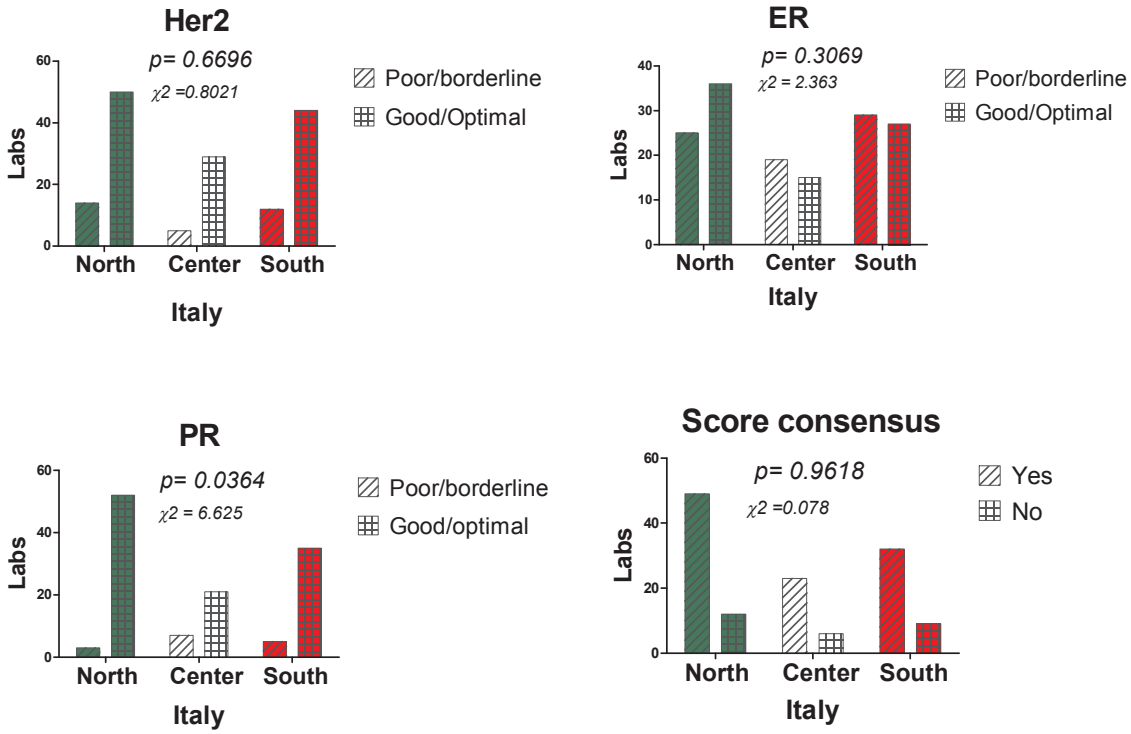
PR							
Concentrated antibodies	**N**	**Vendor**	**O**	**G**	**B**	**P**	**Suff.**
mAb clone **PgR 1294**	5	Dako	3	2	0	0	100%
mAb clone **PgR 636**	22	Dako	16	6	0	0	100%
mAb clone **1A6**	5	Leica	1	2	0	2	60%
mAb clone **16**	13	Leica	11	1	0	1	92%
Ready to use	**N**	**Vendor**	**O**	**G**	**B**	**P**	**Suff.**
mAb clone **16 PA0312**	6	Leica	5	1	0	0	100%
mAb **PgR 636 IR/ISO68**	10	Dako	7	3	0	0	100%
mAb clone **PgR 1294**	1	Dako	0	1	0	0	n.v.
rmAb clone **1E2 790-2223/4296**	60	Ventana	17	31	10	2	80%
mAb clone **PR88** AM328-5 ME	1	Biogenex	1	0	0	0	n.v.
Total	123	61	47	10	5		
Proportion			50%	38%	8%	4%	88%
HER-2							
FDA/CE IVD approved HER2 assays	**N**	**Vendor**	**O**	**G**	**B**	**P**	**Suff.**
PATHWAY, rmAb clone **4B5, 790-2991**	51	Ventana	39	8	0	4	92%
CONFIRM, rmAb clone **4B5, 790-4493**	7	Ventana	6	1	0	0	100%
HercepTest **SK001**	20	Dako	9	8	0	3	85%
HercepTest **K5207**	8	Dako	0	4	1	2	50%
HercepTest **K5204**	7	Dako	0	5	1	1	71%
Oracle mAb clone **CB11, TA9145**	8	Leica	0	2	0	6	25%
Antibodies for laboratory developed HER2 assays, conc. antibody	**N**	**Vendor**	**O**	**G**	**B**	**P**	**Suff.**
mAb clone **CB11**	10	Leica/Cell Marque	1	5	1	3	60%
pAb clone **A0485**	10	Dako	1	7	0	2	80%
Antibodies for laboratory developed HER-2 assays, RTU	**N**	**Vendor**	**O**	**G**	**B**	**P**	**Suff.**
pAb **E2441**	1	Spring Bioscience	0	0	0	1	n.v
Total	122		56	40	3	23	
Proportion			46%	33%	2%	19%	79%

In our study, HER2 scores for the multi-tissue sections of 80-84% of participants were in concordance with the NordiQC assessor group that used ASCO/CAP 2013 interpretation guidelines. The proportion was slightly lower than other previous runs in NordiQC. The greatest discrepancy was recorded for the case with an intermediate staining reaction.

Breast cancer represents worldwide the most frequent cancer with an estimated incidence for women in Italy in 2012 of 31,21% and a mortality rate of 17,05% [22]. Worldwide, an increase of deaths due to this disease is being recorded mainly in less developed countries where the correct management of the disease must deal withpoor financial sources [23]. Sophisticated diagnostic and therapeutic tools are available only in large specialized centers which, however, serve a small portion of patients. Due to this, the 2015 St Gallen International Expert Consensus [13] agreed on prioritizing IHC as a valid surrogate of molecular testing in defining clinically useful subgroups of breast cancer. Four main cancer subgroups may be highlighted: triple negative, ER negative and HER2 positive, ER positive and HER2 positive, and ER positive and HER2 negative (luminal type A and B) [24 25]. Great uncertainty stands in the best treatment of luminal types [26-28]: it strongly depends on the level of ER expression, the certainty of HER2 negativity and the proliferation index. It follows that the accuracy of these immunohistochemical tests allows to address the right therapeutic choices, hence the right management of founds and resources.

Italy is the European country with the largest internal difference of GDP/capita between internal regions [29] with the GDP of the poorest region being only 1/3 of that of Lombardia (the richest). Furthermore, a high variation in the quality and outcomes of care by region is registered [30]. The Euro Health Consumer Index in 2015 evaluated the performance of healthcare provision from a consumer viewpoint. Although the Italian healthcare system is formally under the guidance of one central ministry of health, the national Index score of Italy was

Fig. 7. Laboratories were grouped into three macroareas (North, Center and South). The application of the Fisher's exact test showed that the distribution of HER2, ER and HER2 score consensus results was the same in all three macroareas (respectively p = 0.6696 and χ^2 = 0.8021, p = 0.3069 and χ^2 = 2.363, p = 0.9618 and χ^2 = 0.078); for PR, instead, the proportion of sufficient results was significantly higher in the North than in the South and Center (p = 0.0364 and χ^2 = 6.625) of Italy.



influenced by the optimal performance of regions from the North of Italy and Rome and the poor services offered by most of the regions from the South of Italy. The most alarming datum was the substantial flow across regions of patients seeking better quality care; generally they move from the south to the north of the country [31 32]. What emerges from our study is that no substantial differences exist among the three Italian macro-areas (North, Center and South) in the quality of IHC performed for breast cancer. No statistically significant difference was even found between laboratories that perform more or less than 100 tests/year.

In Italy, each region independently organizes procedures of accreditation to safeguard the quality of healthcare system [33]. A voluntary system of accreditation is also possible whereby it comes out that in regions with higher GDP/capita a woman affected by breast cancer has more chances than in other regions.

Results coming from our project have shown that the overall pass rate of HER2 in RunB19x (79%) was similar to the 86% pass rate in the "official" RunB19. However, these results are significantly lower than those observed by the NordiQC in the previous runs of the same module (see www.nordiqc.org). It is well known that laboratories participating in the same run more than once showan improvement of performance [34]. Quality assessment programs have proven efficient over time and should be a required step in the management of IHC laboratories, especially in the context of predictive markers. The evaluation of the expression of ER and PR receptors and HER2 on tissue is the starting point that initiates a complex process made up of resources (drugs, staff, facilities) and funds. Accurate HER2 testing is essential because of the severe adverse events that may be provoked by HER2-directed therapies in case of their inappropriate administration. Furthermore, cost analyzes have shown that incorrect HER2 tests may also have relevant economic consequences [14].

For this reason the assessment of quality should represent a service offered by the National Healthcare System. Virtuous countries are those where quality certifications are well regulated and mandatory [17]. The effort made by the SIAPEC/IAP with this experience must be seen as the starting point for a project to be carried out and it gave the opportunity to portray the Italian setting that is arranged in a context with several organizational, socio-cultural and geographical diversities compared to other European countries.

References

[1] Taylor CR. *The total test approach to standardization of immunohistochemistry.* Arch Pathol Lab Med 2000;124:945-51.

[2] Torlakovic EE, Nielsen S, Vyberg M, et al. *Getting controls under control: the time is now for immunohistochemistry.* J Clin Pathol 2015;68:879-82.

[3] Goldstein NS, Hewitt SM, Taylor CR, et al. G; Members of Ad-Hoc Committee on Immunohistochemistry Standardization. *Recommendations for improved standardization of immunohistochemistry.* Appl Immunohistochem Mol Morphol 2007;15:124-33.

[4] 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.

[5] Vassilakopoulou M, Parisi F, Siddiqui S, et al. *Preanalytical variables and phosphoepitope expression in FFPE tissue: quantitative epitope assessment after variable cold ischemic time.* Lab Invest 2015;95:334-41.

[6] Hammond ME, Hayes DF, Wolff AC, 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 Oncol Pract 2010;6:195-7.

[7] Wolff AC, Hammond ME, Hicks DG, 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.* Arch Pathol Lab Med 2014;138:241-56.

[8] Nielsen S. *External quality assessment for immunohistochemistry: experiences from NordiQC.* Biotech Histochem 2015;90:331-40.

[9] Vyberg M, Nielsen S. *Proficiency testing in immunohistochemistry-experiences from Nordic Immunohistochemical Quality Control (NordiQC).* Virchows Arch 2016;468:19-29.

[10] Cheung CC, Neufeld H, Lining LA, et al. *The laboratory score/reference method score ratio (LSRSR) is a novel tool for monitoring laboratory performance in immunohistochemistry proficiency testing of hormone receptors in breast cancer: the CIQC experience.* Am J Clin Pathol 2011;136:67-73.

[11] Torlakovic EE, Nielsen S, Francis G, et al. *Standardization of positive controls in diagnostic immunocytochemistry: recommendations from the International Ad Hoc Expert Committee.* Appl Immunohistochem Mol Morph 2015;23:1-18.

[12] Torlakovic EE, Francis G, Garratt J, et al. *Standardization of negative controls in diagnostic immunohistochemistry recommendations from the international ad hoc expert panel.* Appl Immunohistochem Mol Morphol 2014;22:241-52.

[13] Coates AS, Winer EP, Goldhirsch A, et al. *Tailoring therapies-improving the management of early breast cancer: St Gallen International Expert Consensus on the primary therapy of early breast cancer.* Ann Oncol 2015;26:1533-46.

[14] Vyberg M, Nielsen S, Røge R, et al. *Immunohistochemical expression of HER2 in breast cancer: socioeconomic impact of inaccurate tests.* BMC Health Serv Res 2015;15:352.

[15] *The World Health Organization's ranking of the world's health systems.* Photius.com. Retrieved 2015-09-07.

[16] *Country comparison: life expectancy at birth. The World Factbook.* Central Intelligence Agency. 2015. Retrieved 21 December 2015

[17] *Outcomes in EHCI 2015* (PDF). Health Consumer Powerhouse. 26 January 2016

[18] Brügmann A, Eld M, Lelkaitis G, et al. *Digital image analysis of membrane connectivity is a robust measure of HER2 immunostains.* Breast Cancer Res Treat 2012;132:41-9.

[19.] Gown AM. *Current issues in ER and HER2 testing by IHC in breast cancer.* Mod Pathol 2008;21(Suppl 2):S8-S15

[20] Taylor CR. *Predictive biomarkers and companion diagnostics. The future of immunohistochemistry: "in situ proteomics," or just a "stain"?* Appl Immunohistochem Mol Morphol 2014;22:555-61.

[21] Peikari M, Gangeh MJ, Zubovits J, et al. *Triaging diagnostically relevant regions from pathology whole slides of breast cancer: a texture based approach.* IEEE Trans Med Imaging 2016;35:307-15.

[22] http://eco.iarc.fr/eucan/Country.aspx?ISOCountryCd=380.

[23] Ferlay J, Soerjomataram I, Ervik M, et al. *GLOBOCAN 2012 v 1.0, Cancer Incidence and Mortality Worldwide.* 2013. Lyon, International Agency for Research on Cancer. IARC CancerBase No. 11.

[24] Cheang MC, Chia SK, Voduc D, et al. *Ki67 index, HER2 status, and prognosis ofpatients with luminal B breast cancer.* J Natl Cancer Inst 2009;101:736-50.

[25] Prat A, Cheang MC, Martín M, et al. *Prognostic significance of progesteronereceptor-positive tumor cells within immunohistochemically defined luminal A breast cancer.* J Clin Oncol 2013;31:203-9.

[26] 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.* Ann Oncol 2013;24:2206-23.

[27] Coates AS, Millar EK, O'Toole SA, et al. *Prognostic interaction between expressionof p53 and estrogen receptor in patients with node-negative breast cancer: results from IBCSG Trials VIII and IX.* Breast Cancer Res 2012;14:R143.

[28] Regan MM, Pagani O, Walley B, et al. *Premenopausal endocrine-responsiveearly breast cancer: who receives chemotherapy?* Ann Oncol 2008;19:1231-41.

[29] http://ec.europa.eu/eurostat/documents/2995521/6839731/1-21052015-AP-EN.pdf/c3f5f43b-397c-40fd-a0a4-7e68e3bea8cd.

[30] OECD. *Geographic variations in health care: what do we know and what can be done to improve health system performance?* Paris: OECD Publishing 2015.

[31] France G, Taroni F, Donatini A. *The Italian health-care system*. Health Economics 2005;14:187-202.

[32] Ministry of Health. *Sistema di valutazione e monitoraggio della qualità dell'assistenza e delle performance dei sistemi sanitari"*. Relazione sullo stato sanitario del Paese 2009-2010, Ministero della Salute, Roma 2011. Available at www.rssp.salute.gov.it/rssp/paginaParagrafoRssp.jsp?sezione=risposte&capitolo=valutazione&id=2677

[33.] http://www.accredia.it/context.jsp?area=6&ID_LINK=1230

[34.] Wasielewski Rv, Hasselmann S, Rüschoff J, et al. *Proficiency testing of immunohistochemical biomarker assays in breast cancer.* Virchows Arch 2008; 453:537-43