The Breast 30 (2016) 151–155

Contents lists available at ScienceDirect

The Breast

journal homepage: www.elsevier.com/brst

Original article

Discordance in pathology report after central pathology review: Implications for breast cancer adjuvant treatment

Laura Orlando [a, *], Giuseppe Viale [b], Emilio Bria [c], Eufemia Stefania Lutrino [a], Isabella Sperduti [d], Luisa Carbognin [c], Paola Schiavone [a], Annamaria Quaranta [a], Palma Fedele [a], Chiara Caliolo [a], Nicola Calvani [a], Mario Criscuolo [a], Saverio Cinieri [a]

[a] Medical Oncology Division & Breast Unit, 72100 Brindisi, Italy
[b] Pathology Division, European Institute of Oncology, University of Milan, 20141 Milan, Italy
[c] Medical Oncology & Breast Unit, University of Verona, Azienda Ospedaliera Universitaria Integrata, 37134 Verona, Italy
[d] Biostatistics, Regina Elena National Cancer Institute, 00144 Rome, Italy

ARTICLE INFO

Article history:
Received 7 August 2016
Received in revised form
19 September 2016
Accepted 20 September 2016

Keywords:
Early breast cancer
Pathology report

ABSTRACT

Aim: Pathological predictive factors are the most important markers when selecting early breast cancer adjuvant therapy. In randomized clinical trials the variability in pathology report after central pathology review is noteworthy. We evaluated the discordance rate (DR) and inter-rater agreement between local and central histopathological report and the clinical implication on treatment decision.

Methods: A retrospective analysis was conducted in a series of consecutive early breast cancer tumors diagnosed by local pathologists and subsequently reviewed at the Pathology Division of European Institute of Oncology. The inter-rater agreement (k) between local and central pathology was calculated for Ki-67, grading, hormone receptors (ER/PgR) and HER2/neu. The Bland–Altman plots were derived to determine discrepancies in Ki-67, ER and PgR. DR was calculated for ER/PgR and HER2.

Results: From 2007 to 2013, 187 pathology specimens from 10 Cancer Centers were reviewed. Substantial agreement was observed for ER (k0.612; 95% CI, 0538–0.686), PgR (k0.659; 95% CI, 0580–0.737), Ki-67 (k0.609; 95% CI, 0.534–0.684) and grading (k0.669; 95% CI, 0.569–0.769). Moderate agreement was found for HER2 (k0.546; 95% CI, 0444–0.649). DR was 9.5% (negativity to positivity) and 31.7% (positivity to negativity) for HER2 and 26.2% (negativity to positivity) and 12.5% (positivity to negativity) for ER/PgR. According to changes in Her2 and ER/PgR status, 23 (12.2%) and 33 (17.6%) systemic prescription were respectively modified.

Conclusions: In our retrospective analysis, central pathological review has a significant impact in the decision-making process in early breast cancer, as shown in clinical trials. Further studies are warranted to confirm these provocative results.

Introduction

The introduction of adjuvant systemic treatment into early breast cancer management has led to an improvement in overall breast cancer survival. Estrogen receptor (ER), progesterone receptor (PgR) and human epidermal growth factor receptor-2 (HER2) are strong predictors of efficacy of adjuvant therapy in early breast cancer. The magnitude of the impact of endocrine therapy, chemotherapy and targeted therapy is mainly based on hormonal receptor status (HR) and HER2 status in addition to proliferative markers and on tumor grade [1]. Accurate assessment of pathological parameters is mandatory in the-decision making process of systemic therapy in breast cancer patients.

The American Society of Clinical Oncology (ASCO)-College of American Pathologists (CAP) recommended guidelines for both HER2 and ER and PgR immunohistochemical testing, thus producing an algorithm that relies on accurate and reproducible assays [2,3].

In large clinical trials, central pathology review is usually mandatory. In the Breast International Group (BIG) 1-98 trial, central review changed the assessment of HR status in a substantial

* Corresponding author. Medical Oncology Division & Breast Unit, Via Appia, 72100 Brindisi, Italy. Fax: +39 0831537918.
E-mail address: laura.orlando68@gmail.com (L. Orlando).

http://dx.doi.org/10.1016/j.breast.2016.09.015
0960-9776/

proportion of patients [4]. Of 6100 women classified ER positive in local assessment, central review found 66 ER negative (1.1%) and 54 low ER (0,9%). The discordance was more marked for PgR. In the ALTTO trial for HER2 positive disease, HER2 and ER were centrally reviewed by the Mayo Clinic in Rochester and the European Institute of Oncology in Milan (IEO). Among locally HER2 positive tumors, 5.8% and 14.5% were centrally negative for the Mayo and the IEO respectively. Among locally ER positive tumors, 16.2% and 4.2% were found negative at the Mayo and the IEO central review respectively [5]. For other pathological parameters, such as Ki67 and grading, the rate of discordance rate appears more marked [6].

Despite the multiple data of discordance rate after central review in breast cancer, the potential clinical impact outside clinical trials remains limited. Previous studies of inter-institutional pathology consultations for breast cancer reported a 4–29% discordance rate, however, information on specific discordant parameters is limited [7,8].

The present study reports the results of the central pathology review of ER, PgR, HER2 status, Ki67 and grading of early breast cancer and the implications for the selection of adjuvant systemic therapies.

Materials and methods

We conducted a retrospective review of 210 consecutive invasive breast cancer specimens referred to our Institution from 2007 to 2013. Specimens were sent for central pathological review to the European Institute of Oncology (IEO) in Milan. One hundred eighty seven samples were selected for this analysis.

Local HER2, ER, PgR, Ki67 and grading refer to the initial testing performed on the tumor tissue samples. Central HER2, ER, PgR, Ki67 and grading refer to the results from the IEO review.

The invasive component was confirmed in all specimens. Two tumors were excluded from the analysis due to the presence of advanced disease, 13 because only the primary core biopsy was available, and 8 because only hormonal receptor review was performed. The medical records of patients who had discordant diagnoses were reviewed in order to evaluate changes in the management plan.

The study was approved by the Local Ethical Committee.

Pathology

All the pathology reviews were performed at the IEO (Milan). The same assay and methodology were applied to each sample. The central laboratory run the analysis on the same paraffin block used in local laboratories.

IHC and FISH for HER2 were performed using the HercepTest® kit (Dako, Glostrup, Denmark) and PathVysion HER2 DNA probe kit/HER2/centromere 17 probe mixture (Abbott Molecular, Des Plaines, IL). HER2 positivity was defined according to the FDA scoring system, (intense circumferential membrane staining in >10% of tumor cells by IHC or HER2 gene copy number/CEP17 signals $\geq$ 2 by FISH).

IHC for ER and PgR was tested centrally using the DAKO ER/PR PharmDX kit, and defined positive if $\geq$ 1% immunostained tumor cells [1].

Statistical analysis

Licenced MedCalc (v. 11.0) was used to analyze the inter-rater variability between the local pathological diagnosis and the central review, according to the Kappa (k) index. The index was interpreted according to the following values: <0.20 (bad); 0.21–0.40 (poor); 0.41–0.60 (moderate); 0.61–0.80 (good); and 0.81–1.00 (excellent) [9]. The significance level (p) was taken as 0.05.

In order to visually test and weigh differences between local and central pathology, the Bland–Altman plots were determined for Ki-67, ER and PgR [10]. Results obtained by central pathology review (retesting) were compared with local tested results and the discordance rate (DR) and inter-rater agreement were calculated. Tumors with one or more target parameters that were unknown or missed (ER, PgR, HER2, Ki67, histologic type, grading, Ki67) were excluded. Correlation analysis between local and central pathology was also conducted l for ER, PgR, and Ki67, according to parametric (Pearson's r, with 95% confidence intervals, CI) and non-parametric (Spearman's Rho and Kendall's Tau) coefficients; a regression equation was calculated according to the regression analysis (parametric R^2) [11]. DR was defined as the positive-to-negative or negative-to-positive changes according to ER/PgR status or according to HER2 status. Any main changes in treatment decision from initial purpose to final prescription were also considered: addition or subtraction of endocrine therapy and/or of anti HER-2 therapy. These main changes were calculated as percentage.

Results

A total of 210 specimens of invasive breast cancer from ten Cancer Centers were reviewed. 23 specimens were excluded from the analysis: two because of the presence of metastatic disease at diagnosis, 13 for whom only biopsy samples were available, 8 because they were re-tested only for hormonal receptors. Median age of patients was 52 years (28–76), 100 (53.4%) patients were postmenopausal (Table 1.)

Local analysis revealed 145 tumors as ER- and/or PgR-positive (77.5%) and 41 tumors as HER2-positive (21.9%). At central review 136 (72.7%) and 42 (22.4%) tumors were ER- and/or PgR-positive and HER2-positive respectively.

Substantial agreement was observed for ER (Kappa = 0.612; 95% CI, 0538–0.686), PgR (Kappa = 0.659; 95% CI, 0580–0.737), Ki67 (Kappa = 0.609; 95% CI, 0.534–0.684) and grading (Kappa = 0.669; 95% CI, 0.569–0.769). Moderate agreement was found for HER2 (Kappa = 0.546; 95% CI, 0444–0.649) (Table 2). The analysis confirmed the dispersions of values according to ER, PgR and Ki67 (Figs. 2–3). The Bland–Altman plot did confirm the absence of major differences or discrepancies between the two assays for ER, PgR and Ki67 (Figs. 1–3). Supplementary Fig. S1 shows the correlation between the two pathologic determinations for the same variables. With regard to HER2 distribution, detailed descriptors are reported in Supplementary Fig. A2.

Table 1
Patients and tumors features at diagnosis.

	n (%)
Age (median) (y)	52 (28–76)
Pre/postmenopausal	87/100
pT1	62 (33%)
pT2	56 (30%)
pT3	30 (16%)
pT4	39 (20.8%)
pN1	102 (54.5%)
pN2	65 (34.7%)
pN3	20 (10.7%)
Histology	
Ductal	158 (84.4%)
Lobular	17 (9%)
Others	12 (6.4%)
ER and/or PgR pos	145 (77.5)
ER and PgR neg	42 (22.5)
HER2 positive	41 (21.9)
Triple negative	25 (13.3%)

HER2 positive: staining 3+. *ER and/or PgR positive*: staining $\geq$ 1%.

Table 2
Inter-rater agreement (Kappa) between Local and Central Review of Estrogen/Progesterone receptors (ER/PgR), Ki67 expression, Grading and HER2 immunohistochemistry.

	Weighted Kappa	Standard error	95% C.I.
ER	0.612 [Linear Weights]	0.038	0.538–0.686
	0.740 [Quadratic Weights]	0.038	0.665–0.815
PgR	0.659 [Linear Weights]	0.040	0.580–0.737
	0.749 [Quadratic Weights]	0.043	0.666–0.833
Ki67	0.609 [Linear Weights]	0.038	0.534–0.684
	0.741 [Quadratic Weights]	0.046	0.651–0.832
Grading	0.669 [Linear Weights]	0.051	0.569–0.769
	0.682 [Quadratic Weights]	0.069	0.546–0.817
HER2	0.546 [Linear Weights]	0.052	0.444–0.649
	0.625 [Quadratic Weights]	0.055	0.517–0.733

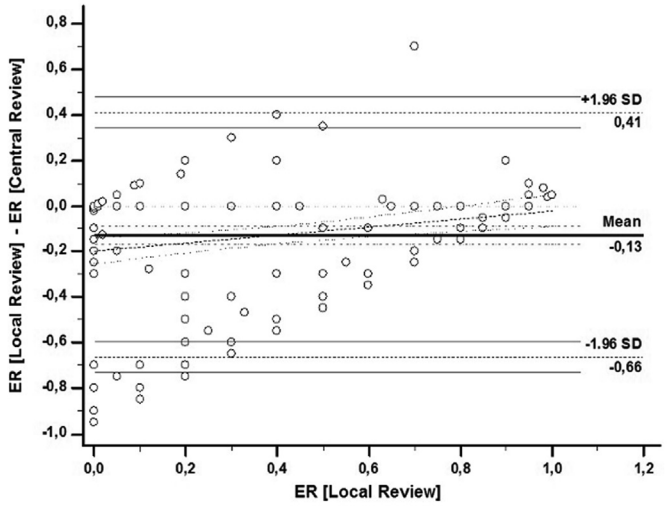

Fig. 1. Bland–Altman plot comparing local and central review of Estrogen Receptor (ER) status expression [differences plotted against ER local review].

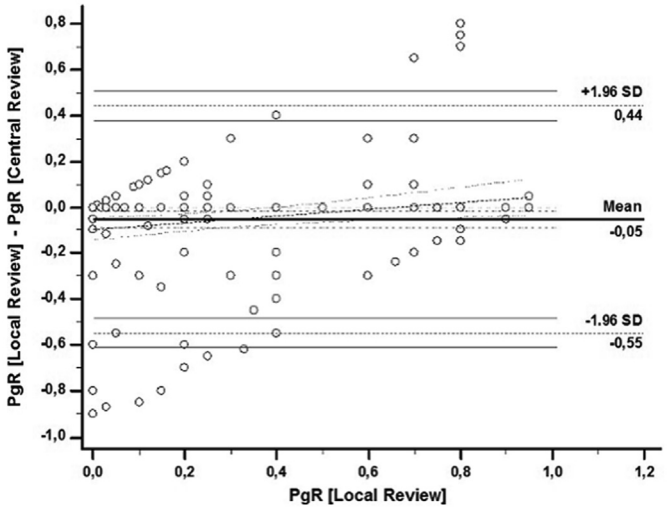

Fig. 2. Bland–Altman plot comparing local and central review of Progesterone Receptor (PgR) status [differences plotted against PgR local review].

Discordance rate: HER2 and ER/PgR

Twelve of 126 locally HER2 0 or 1+ tumors were found positive at central review (11 tumors positive by IHC and one with gene amplification at FISH analysis) (DR 9.5%). Thirteen of 41 locally HER2 3+ were centrally HER2 negative (0 or 1+) (DR 31.7%). Of 20 locally HER +2 tumors, 15 were not at central analysis (DR 75%): 12 were negative (0 or +1, with FISH negative) while 3 were +3 (with FISH positive) (Table 3).

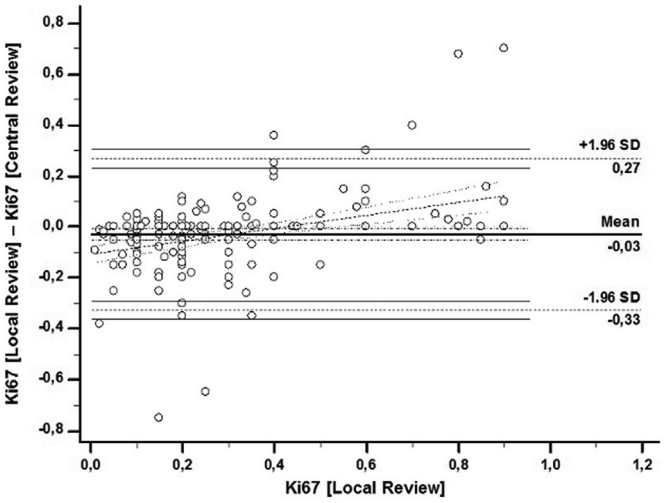

Fig. 3. Bland–Altman plot comparing local and central review of Ki67 expression [differences plotted against Ki67 local review].

Eleven of 42 tumors locally ER- and PgR-negative, were found to be positive at central evolution (DR 26.2%). 18 of 145 locally ER- and PgR-positive specimens had no central staining for both ER and PgR (DR 12.5%) (Table 3).

In our series, after changes in HER2 status and hormonal receptors status, (negativity vs positivity and *vice versa*) 23 (12.2%) and 33 (17.6%) systemic prescriptions were modified respectively.

Discussion

Accurate evaluation and measurement of HR and HER2 status are of clinical importance when selecting adjuvant therapy for early breast cancer.

The present study assessed the DR and inter-rater variability of local and central pathology and its implications in systemic adjuvant treatment for early breast cancer.

By using the Kappa index, we found a substantial agreement for ER and PgR, meaning that values from central review did not extensively differentiate from local data, and a moderate agreement for HER2 evaluation.

However, when we considered DR between local and central analysis, we found a great discordance for HER2 assessment: 12 negative tumors were found positive at central review (11 tumors 3+ and one 2+ with gene amplification at FISH analysis), with a discordance rate of 9.5%, and 13 positive tumors were centrally HER2 negative (0 or 1+), with a discordance rate of 31.7%.

Table 3
Discordance in ER/PgR and HER2 status between local and central laboratories.

Local	Central	n	DR
ER/PgR negative	ER/PgR positive	11	26.2
ER/PgR positive	ER/PgR negative	18	12.5
HER2 positive	HER2 negative	13	31.7
HER2 negative	HER2 positive	12	9.5

ER estrogen receptor, *PgR* progesterone receptor, *HER2* human epidermal growth factor 2.
IHC immunohistochemical, *FISH* fluorescent in situ hybridization.
DR discordance rate (%).

Regarding HR status, 11 of 42 tumors which were locally ER- and PgR negative, resulted positive at central evaluation, while 18 of 145 locally ER- and PgR-positive tumors were found negative at central review (DR 26.2% and 12.5% respectively).

These results are worse than those previously reported, mainly in terms of local HER2-positive results found negative at central review, and HR-negative results found positive after central review [8].

The DR is likely a better parameter when clinical impact of the results is considered. Changes from positivity to negativity for HR and HER2 have implications for adding or omitting endocrine or anti-HER2 therapies. In accordance with the policy of our center, tumors with $\geq$1% ER or PgR immunostained tumor cells receive endocrine therapy (with or without chemotherapy) and which are HER2 positive tumors receive trastuzumab. In our series, systemic prescriptions were modified in 12.2% and 33% of the cases according to changes in HR and HER2 status respectively.

We have not evaluated the changes performed in chemotherapy prescription according to central pathology review because the choice of adding chemotherapy to HR-positive tumors is complex, depending upon both immunohistochemical and clinical parameters. However, the distribution of ER values reported by the Kappa index (Fig. 1) could have an impact on chemotherapy indication. The ER values distribution might also play a role in predicting the benefit of endocrine therapy. Recently San Gallen endocrine response classes have been used to predict recurrence rates over time and to demonstrate a marked variability in endocrine therapy benefit according to the level of ER positivity [12].

In our study we have analyzed ER and PR together because the role of PgR expression is not completely clarified in terms of prognostic and predictive significance in early breast cancer [13], although recently the prognostic impact of PgR loss has been found associated with poor prognosis in Luminal B breast tumors [14].

In our analysis, the inter-rater agreement according to Ki67 values was substantial (Kappa value: 0.612), which means consistency between local and central laboratories. Usually, greater variability in Ki67 values is observed in different laboratories, due to interlaboratory differences in staining methodology, scoring interpretation, including cut-off determination [6]. Also histological grading is of clinical relevance in breast cancer, because it can contribute to selecting the adjuvant therapy [1]. However, it can be subject to operator-dependent variability. In our series, a greater agreement was detected compared to other parameters analyzed. (Kappa value: 0.699). This data could be justified by observing that the majority of tumors were locally grade 3 and that the definition of undifferentiated tumors is usually easier than that for grade 2 tumors and is less subject to inter-operator variability.

Although the finding of discordant values in histopathological parameters might be a consequence of intratumoral heterogeneity, different sections of one single blocks are usually representative of the tumor.

Our findings underlined the need for accurate and standardized analysis of biological parameters for an adequate selection of adjuvant systemic treatment. Different studies have underscored the importance of internal and external quality control for immunohistochemical factors [15].

Immunohistochemistry data can be affected by preanalytic and analytic variability. Preanalytic factors include fixation and processing, both subject to suffering delay, different duration of procedure and type of reagents.

Analytic factors include intra and inter-observer variability in data interpretation and cut points, mainly regarding percentage of cells stained and intensity of staining [16].

We analyzed 210 early breast cancer cases referred as outpatient visits to our center for adjuvant treatment planning. In consideration of the retrospective nature of our work, the likelihood of selection bias cannot be excluded. Pathology review is not a standard procedure, so the decision to performing it might derive from tumor or patient features or from the knowledge of a specific source of pathology sample. However the majority of cases (65%) derived from a single institution, which used standardized preanalytical and analytical methods.

The IHC evaluation of ER, PgR and HER2 status may be considered the final expression of molecular profiles. Thus, in order to overcome the potential weakness of IHC reproducibility and provide quantitative measurement of the therapeutic target, assays based on mRNA analysis have been developed [17] and compared with IHC assays. Results are controversial and do not support the use of molecular profiles as a substitute for IHC techniques. In ECOG 2197 trial, a case-control evolution was conducted to compare ER and PgR status by local IHC, central ICH and central-reverse-transcriptase polymerase chain reaction (RT-PCR) by using the 21-gene assay [18]. The authors found a high degree of concordance among local IHC, central IHC and gene assay. Park and colleagues found similar concordance between IHC evaluation of ER and PgR and RT-PCR assay (Oncotype DX), but showed poorer agreement between FISH HER2 amplified and oncotype DX. They discouraged the preferential use of Oncotype DX over IHC and FISH evaluation of HER2 [19]. In contrast to these results, a great concordance between central FISH and quantitative RT-PCR using Oncotype Dx for HER2 status was reported (97%, 95% CI, 96%–99%) in a large case-control study [20]. More recently, an unacceptable false negative rate was reported in an independent quality assurance study by using Oncotype Dx test [21]. Up to now, there are therefore controversial and insufficient data to support the use of quantitative RT-PCR as a substitute for IHC or fluorescent in situ hybridization assays, and much efforts should be dedicated toward improving the reproducibility of ICH.

Finally, the economic impact of a second revision of the original pathologic material cannot be underestimated, since this is not something currently covered by the Italian National Health Service. However, the increasing likelihood of exact diagnosis after review by specialized breast pathologists may lead to a significant impact on patient care and outcome.

In summary, the variability in IHC results reported in our series confirms the data already shown in other work. However, these data reflect a local situation and do not suggest the need to centrally re-test every tumor.

The inaccuracy of immunohistochemistry might lead to inappropriate decisions about adjuvant strategies in early breast cancer and should be avoided. Efforts should be made to reduce the discordance rate between local and central laboratories. Expert and skilled pathologists, as well as quality control programs are desirable in order to improve the management of early breast cancer.

Conflict of interest statement

Authors declare no conflict of interest.

Acknowledgment

We thank Dr. Cosimo Caforio for excellent technical assistance.

Appendix A. Supplementary data

Supplementary data related to this article can be found at http://dx.doi.org/10.1016/j.breast.2016.09.015.

References

[1] Coates AS, Winer EP, Goldhirsch A, Gelber RD, Gnant M, Piccart-Gebhart M, 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.

[2] Wolff AC, Hammond ME, Hicks DG, Dowsett M, McShane LM, Allison KH, et al. Recommendations for human epidermal growth factor receptor 2 testing in breast cancer: American Society of Clinical Oncology/College of American Pathologists clinical practice guideline update. J Clin Oncol 2013;31:3997–4013.

[3] Hammond MEH, Hayes DF, Dowsett M, Allred DC, Hagerty KL, Badve S, et al. American Society of Clinical Oncology/College of American Pathologists guidelines recommendations for immunohistochemical testing of estrogen and progesterone receptors in breast cancer. J Clin Oncol 2010;28:2784–95.

[4] Viale G, Regan MM, Maiorano E. Prognostic and predictive value of centrally reviewed expression of estrogen and progesterone receptors in a randomized trial comparing letrozole and tamoxifen adjuvant therapy for postmenopausal early breast cancer: BIG 1-98. J Clin Oncol 2007;25:3846–52.

[5] McCullough AE, Dell'Orto P, Reinholz MM, Gelber RD, Dueck AC, Russo L, et al. Central pathology laboratory review of HER2 and ER in early breast cancer: an ALTTO trial [BIG 2-06/NCCTG N063D (Alliance)] Ring study. Breast Cancer Res Treat 2014;143(3):485–92.

[6] Polley MYC, Leung SCY, McShane LM, Gao D, Hugh JC, Mastropasqua MG, et al. An international Ki67 reproducibility study. J Natl Cancer Inst 2013;105:1897–906.

[7] Chang JH, Vines E, Bertsch H, Fraker DL, Czerniecki BJ, Rosato EF, et al. The impact of a multidisciplinary breast cancer centre on recommendations for patient management: the University of Pennsylvania experience. Cancer 2001;91:1231–7.

[8] Price JA, Grunfeld E, Barnes PJ, Rheaume DE, Rayson D. Inter-institutional pathology consultations for breast cancer: impact on clinical oncology therapy recommendations. Current Oncol 2010;17:25–32.

[9] Landis JR, Koch GG. The measurement of observer agreement for categorical data. Biometrics 1977;33(1):159–74.

[10] Bland JM, Altman DG. Measuring agreement in method comparison studies. Stat Methods Med Res 1999;8:135–60.

[11] Bria E, Massari F, Maines F, Pilotto S, Bonomi M, Porta C, et al. Progression-free survival as primary endpoint in randomized clinical trials of targeted agents for advanced renal cell carcinoma. Correlation with overall survival, benchmarking and power analysis. Crit Rev Oncol Hematol 2015;93(1):50–9.

[12] Koornstra RHT, Beelen KJ, Vincent AD, van der Noort V, van Diest PJ, Linn SC. St Gallen endocrine response classes predict recurrence rates over time. Breast 2015;24(6):705–12.

[13] Dowsett M, Ebbs SR, Dixon JM, Skene A, Griffith C, Boeddinghaus I, et al. Biomarker changes during neoadjuvant anastrozole, tamoxifen, or the combination: influence of hormonal status and HER 2 in breast cancer – a study from IMPACT trialists. J Clin Oncol 2005;23:2477–92.

[14] Cancello G, Maissonneuve P, Rotmens N, Viale G, Mastropasqua MG, Pruneri G, et al. Progesterone receptor loss identifies Luminal B breast cancer subgroups at higher risk of relapse. Ann Oncol 2013;24:661–8.

[15] Kirkegaard T, Edwards J, Tovey S, McGlynn LM, Krishna SN, Mukherjee R, et al. Observer variation in immunohistochemical analysis of protein expression, time for change? Histopathology 2006;48:787–94.

[16] Rodhes A, Jasani B, Balaton AJ, Miller KD. Immunohistochemical demonstration of oestrogen and progesterone receptors: correlation of standards achieved on in house tumours with that achieved on external quality assessment material in over 150 laboratories from 26 countries. J Clin Pathol 2000;53:292–301.

[17] Kim C, Paik S. Gene-expression-based prognostic assays for breast cancer. Nat Rev Clin Oncol 2010;7:340–7.

[18] Badve SS, Baehner FL, Gray RP, Childs BH, Maddala T, Liu ML, et al. Estrogen and progesterone-receptor status in ECOG 2197: comparison of immunohistochemistry by local and central laboratories and quantitative reverse transcription polymerase chain reaction by central laboratory. J Clin Oncol 2008;26:2473–81.

[19] Park MM, Ebel JJ, Zhao W, Zynger DL. ER and PgR immunohistochemistry and HER2 FISH versus oncotype DX: implications for breast cancer treatment. Breast J 2014;20:37–45.

[20] Baehner FL, Achacoso N, Maddala T, Shak S, Quesenberry Jr CP, Goldstein LC, et al. Human epidermal growth factor receptor 2 assessment in a case-control study: comparison of fluorescent in situ hybridization and quantitative reverse transcription polymerase chain reaction performed by central laboratories. J Clin Oncol 2010;28:4300–6.

[21] Dabbs DJ, Klein ME, Mohsin SK, Tubbs RR, Shuai Y, Bhargava R. High-false negative rate of HER2 quantitative reverse transcription polymerase chain reaction of the oncotype Dx test: an independent quality assurance study. J Clin Oncol 2011;29:4279–85.