A MULTIDISCIPLINARY REVIEW OF THE IMAGING AND FUNCTIONAL RESPONSE TO NEOADJUVANT TREATMENT OF EARLY-STAGE BREAST CANCER

Ana María García-Vicente, Myriam Montes Fernandez, Ricardo Sánchez-Escribano, Ana Paula Caresia-Aróztegui, Sara Jiménez Arranz, Pedro Sánchez Rovira, Aurora Crespo de la Jara, Francisco Javier de Castro-García, José E. Alés-Martínez

Edición: Sociedad Española de Oncología Médica (SEOM).
Sociedad Española de Medicina Nuclear e Imagen Molecular (SEMNIM), Sociedad Española de Radiología Médica (SERAM)
Diseño y maquetación: Gonext Producciones S.L.
ISBN: 978-84-09-36932-4

A MULTIDISCIPLINARY REVIEW OF THE IMAGING AND FUNCTIONAL RESPONSE TO NEOADJUVANT TREATMENT OF EARLY-STAGE BREAST CANCER

Ana María García-Vicente[a,j*], Myriam Montes Fernandez[b,k,*], Ricardo Sánchez-Escribano[c,l,*], Ana Paula Caresia-Aróztegui[d,j,*], Sara Jiménez Arranz[e,k,*], Pedro Sánchez Rovira[f,l,*], Aurora Crespo de la Jara[g,j,*], Francisco Javier de Castro-García[h,k,*], José E. Alés-Martínez[i,l,*]

[a]Nuclear Medicine Department, University General Hospital of Ciudad Real, Calle Obispo Rafael Torija, s/n, 13005 Ciudad Real, Spain, E-mail: angarvice@yahoo.es / [b]Radiology Department, Hospital Clínico San Carlos, Calle Profesor Martín Lagos, s/n, 28040 Madrid, Spain, E-mail: myriammfag@hotmail.com / [c]Medical Oncology Department, Hospital Clínico Universitario de Valladolid (HCUV), Avenida Ramón y Cajal, s/n, 47003 Valladolid, Spain, E-mail: rsescribano@seom.org / [d]Nuclear Medicine Department, Parc Taulí Hospital Universitari, Calle Parc Taulí, 1, 08208 Sabadell, Spain, E-mail: apcaresia@tauli.cat / [e]Radiology Department, Hospital Universitario 12 de Octubre, Avenida de Córdoba, s/n, 28041 Madrid, Spain, E-mail: sarajimenezarranz@gmail.com / [f]Medical Oncology Department, Hospital Universitario de Jaén, Avenida del Ejército Español, 10, 23007 Jaén, Spain., E-mail: oncopsr@yahoo.es / [g]Nuclear Medicine Department, University Hospital of San Juan, Carretera N-332, s/n, 03550 San Juan de Alicante, Spain, E-mail: crespo_aur@gva.es / [h]Radiology Department, Hospital Nuestra Señora de Sonsoles, Avenida Juan Carlos I, s/n, 05004 Ávila, Spain, E-mail: fcojav7769@gmail.com / [i]Medical Oncology Department, Hospital Nuestra Señora de Sonsoles, Avenida Juan Carlos I, s/n, 05004 Ávila, Spain, E-mail: jealesm@seom.org / [j]Spanish Society of Nuclear Medicine and Molecular Imaging (SEMNIM) / [k]Spanish Society of Medical Radiology (SERAM) / [l]Spanish Society of Medical Oncology (SEOM)

*All authors have contributed equally to the content of this manuscript

ABSTRACT

Neoadjuvant treatment of breast cancer consists of the administration of systemic therapies for patients with resectable tumors (stages I-III). The goals are to reduce tumor size to allow for a more conservative resection and to gain early information about the tumor sensitivity to the treatment. Randomized clinical trials have shown that presurgical and postsurgical chemotherapy are in general equivalent in terms of relapse reduction and overall survival. It is of the utmost importance to have imaging methods that can predict pathologic complete response at baseline, permit detection of early responses and correlate with pathological response with sufficient accuracy. In this document, a group of experts from the Spanish Society of Medical Radiology (SERAM), the Spanish Society of Nuclear Medicine and Molecular Imaging (SEMNIM) and the Spanish Society of Medical Oncology (SEOM), provide an update on the evidence for the different available imaging techniques in the assessment of the response to the neoadjuvant treatment of breast cancer under a multidisciplinary approach and make recommendations on how to best provide a multidisciplinary approach in the imaging evaluation of response to neoadjuvant treatment of early breast cancer.

KEYWORDS: breast cancer, neoadjuvant treatment, imaging techniques, pathological response, radiological response

KEY POINTS

- Imaging and functional tumor changes after neoadjuvant treatment of early-stage breast cancer.
- Primary and interim evaluations and prediction of response.
- PET-CT and MRI to predict and evaluate responses to neoadjuvant treatment.
- Intrinsic breast cancer subtypes and differential response to treatment.

BACKGROUND

The neoadjuvant treatment of breast cancer (NAT) consists of the administration of systemic therapies for patients with resectable tumors (stages I-III). The goals are to obtain a reduction of tumor size for a more conservative resection (e.g., converting a mastectomy into a lumpectomy) and to gain early information about tumor treatment sensitivity. Randomized clinical trials have shown that presurgical and postsurgical chemotherapies are equivalent in terms of relapse reduction and overall survival [1].

NAT should be indicated by a multidisciplinary team (including surgeons, oncologists, imaging specialists, and pathologists), considering possible cosmetic results and the probability of response based on tumor type, anatomical conditions, and comorbidities.

NAT can also provide the following additional advantages:

- Early initiation of potentially curative therapies without waiting for postoperative recovery, which can be especially important for fast-growing tumors.
- In vivo evaluation of the efficacy of a therapy in individual patients without compromising the risk of recurrence.
- Pathologic complete response (pCR) can be used as a surrogate marker of survival benefit, optimizing the development of new drugs as well as the design of more effective clinical trials with fewer requirements in terms of follow up and number of patients [2]. The most common method used by pathologists to evaluate pCR is determination of the Residual Cancer Burden (RCB) by a standardized pathological assessment method based on the largest area and cellularity of residual invasive primary cancer and the number of involved lymph nodes and size of largest metastasis. The classification ranges from RCB class 0 (ypT0/isypN0) to RCB class III (extensive residual disease). This classification has demonstrated prognostic significance for long-term survival in all subtypes of breast cancer [3].

Breast cancer is a heterogeneous disease, and various subtypes have been defined using the expression of proteins detectable by immunohistochemistry and gene expression profiles via RNA analysis. These different phenotypes vary in terms of their responses to systemic treatments and require different approaches when evaluating treatment response. Below, we very briefly describe these phenotypes:

Luminal breast cancer

This type is defined by the expression of hormone receptors (HRs), estrogen (ER), and/or progesterone (PR) and includes subtypes A and B. **Luminal A** tumors are characterized by a more favorable clinical course, strong expression of hormone receptors and low proliferation index. They show the best rates of long-term survival and the lowest risk of relapse among all other phenotypes. However, they show a low sensitivity to chemotherapy and have the lowest rates of pCR to NAT. In addition, no correlation has been found between pCR and survival in these patients [4]. **Luminal B** tumors may have lower hormone receptor expression, absence of progesterone receptor expression and/or a higher proliferation index. They have a more aggressive disease course but show greater benefit from chemotherapy, with a better overall pCR-survival correlation.

CORRESPONDENCE
José E. Alés-Martínez, MD
Medical Oncology Department, Hospital Nuestra Señora de Sonsoles, Avda Juan Carlos I, s/n. 05004 Ávila, Spain / Tel: +34619287189 / Email: jealesm@seom.org

Triple-negative breast cancer (TN)
This subtype is also a heterogeneous group (4 to 8 profiles depending on the molecular assays applied, including luminal with androgen receptor expression, basal like, and mesenchymal), with the common finding of lack of expression of HRs and non-amplified c-erbB2. Within this subtype, the **basal phenotype**, as defined by RNA expression profiling, is highly sensitive to chemotherapy, with pCR rates of approximately 50% and a high correlation with survival. However, other subtypes in this group, such as basal 2 or luminal with androgen expression (LAR), show much poorer responses [5].

HER2-positive breast cancer
This subtype is defined by the amplification of the c-erbB2 oncogene and subsequent overexpression of the HER2 protein as the main driver of its pathogenicity. The subgroup not expressing HRs has the highest rate of pCR using double-blockade anti-HER2 therapy (trastuzumab-pertuzumab) and chemotherapy, reaching a rate of over 60% [6, 7]. The correlation between pCR and survival is high, and the Food and Drug Administration (FDA) approved the use of preoperative pertuzumab-trastuzumab based on these results. However, the rate of pCR and its correlation with survival are lower for patients who coexpress HRs (sometimes described as Luminal B-HER2).

In summary, it is of the utmost importance to have imaging methods that can predict pCR at baseline, allow for the detection of early responses and correlate with pathological response with sufficient accuracy. In this document, a group of experts from the Spanish Society of Medical Oncology (SEOM), the Spanish Society of Medical Radiology (SERAM) and the Spanish Society of Nuclear Medicine and Molecular Imaging (SEMNIM) provide an update of the evidence for the different available imaging techniques in the assessment of responses to NAT under a multidisciplinary approach and make a number of practical recommendations to optimize this approach.

CONSENSUS RECOMMENDATIONS FOR IMAGING EVALUATION OF RESPONSE TO NAT IN BREAST CANCER

QUESTION 1: Initial breast imaging

Which is the ideal breast imaging method prior to NAT?

Consensus recommendation
MRI with functional assessment (DCE) should be performed in all patients eligible for NAT as a baseline study prior to therapy; The main value of PET/CT in this setting would be a more comprehensive N and M staging, as long as it could become a more available technology".

Literature review and clinical interpretation
Using magnetic resonance imaging (MRI), some features, such as round or oval shape and the absence of intratumoral T2 high signal, suggest tumor aggressiveness and may be associated with better responses to NAT, whereas an irregular or spiculated mass shape and intratumoral T2 high signal were more frequently observed in nonresponder patients. According to some authors, the presence of peritumoral edema is associated with a worse outcome [8, 9]. Recent studies have described that a combined texture analysis of intratumoral and peritumoral regions from a pretreatment Dynamic Contrast Enhanced-MRI (DCE-MRI) could predict pCR to NAC [10]. Larger prospective studies are needed, but in-house clinical-radiology collaboration is encouraged.

The molecular characteristics of the tumor are fundamental determinants of the final histological response after NAT. Thus, biologically more aggressive tumors are associated with a better response to NAT [11, 12]. This feature explains the greater glycolytic metabolism detected using 18F-fluorodeoxyglucose positron emission tomography/computed tomography (FDG PET/CT) in responding tumors found in previous research [12-14]. However, there is not a specific standardized uptake value (SUV) cutoff to predict response, and other studies have not found such an association [15-17]. The need for additional studies that consider tumor molecular information is justified. However, the molecular heterogeneity of breast cancer translates into variations in the distribution of the uptake intensities in FDG PET/CT that can be quantified by a texture analysis [18-20]. Thus, certain texture characteristics might be associated with a better response to NAT [21], although information is limited [8-10]. Close information exchange among image and clinical specialists is highly recommended.

QUESTION 2: interim evaluation

Is an interim imaging assessment recommended in NAT?
Consensus recommendation
An interim imaging test is advisable for the optimal evaluation and prediction of response to NAT. An effective evaluation of the early response could avoid the toxicity associated with ineffective treatments and help make an informed decision to change the therapeutic strategy for nonresponsive patients.

Which are the preferred methods for the interim assessment of response?
Consensus recommendation
Regarding the optimal method, FDG PET/CT seems to perform better in the early prediction after one to three cycles and MRI in later evaluations (after three or more cycles of NAT). However, diagnostic techniques (RMN vs PET-CT) should be selected based on the availability and diagnostic experience of the imaging team at each site in close collaboration with the clinical team.

Literature review and clinical interpretation
Given the duration of typical NAT regimens, intermediate controls are usually performed to assess early tumor response and to differentiate responders from non-responders after a few cycles of NAT. Functional data extracted from MRI after one or two cycles are useful to differentiate faster responders from non-responder patients. In a systematic review of 13 studies, Marinovich et al. [22, 23] evaluated the accuracy of DCE-MRI to correctly identify non-responders after one or two cycles of NAT. The authors showed that the sensitivity and specificity were greater in the studies that used tumor volume and dynamic sequence quantitative parameters (i.e., reductions in K-trans and early contrast uptake) to evaluate responses but were lower if the tumor was measured in one or two dimensions [23]. Similarly, the ACRIN 6657/I-spy trial showed that volumetric measurements of the tumor (better than measurements of the longest diameter) were the best predictor of early response during treatment [24]. Padhani et al. [25] also showed that K-trans values decreased in responding patients after one or two cycles, as did the size of the tumor, and with a similar level of precision. In addition, the apparent diffusion coefficient (ADC) associated with diffusion-weighted MRI (DWI-MRI) can be used as an early response biomarker because changes in ADC occur before changes in size, enabling us to differentiate responsive patients (who present higher ADC values in their tumors) from nonresponsive patients. The accuracy by which MRI estimates residual lesions after two cycles of NAT depends on the tumor subtype. This technique is more accurate in the TN and HER2+ subtypes as well as for high-grade tumors [26].

The tumor changes that occur during NAT determine the modifications in the metabolic variables obtained on FDG PET/CT (Figure 1 PET). Semiquantitative analysis of these variations is performed by comparing the baseline with interim FDG PET/CT (after 1-3 cycles) or after the end of NAT (final FDG PET/CT) [27] (Figure 2 PET). Normally, the most used metabolic variable is the SUV_{max} (value of the voxel with the highest SUV), although others have used SUV_{mean} (mean of SUV voxels within the tumor volume), SUV_{peak} (mean of SUV voxels in a spherical region of interest of 1 cm^3 near the SUV_{max} voxel), MTV (metabolic tumor volume, determined as the tumor volume with significant FDG uptake), TLG (total lesion glycolysis, determined as the MTV multiplied by the SUV_{mean}), or even texture variables [28, 29]. Calculation of the percentage decrease (Δ%) or reduction rate (RR) is obtained using the following formula: (PET response value - baseline PET value)/baseline PET value X 100.

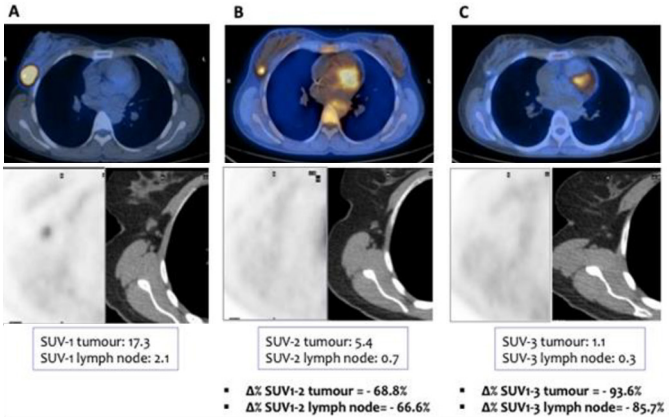

Figure 1. Methodology for the neoadjuvant treatment response, interim and final. (a) Baseline 18F-FDG PET/CT (PET-1), co-registered PET/CT axial slice of mammary lesion, PET and CT of axial slice of axillary lymph node. (b) Interim 18F-FDG PET/TC after 2 cycles of chemotherapy (PET-2) and (c) final treatment 18F-FDG PET/TC (PET-3). In all cases, SUVmax is defined in each location (breast lesion and lymph node), and the calculation of Δ% SUVmax is made between the interim and final treatments with respect to basal PET.

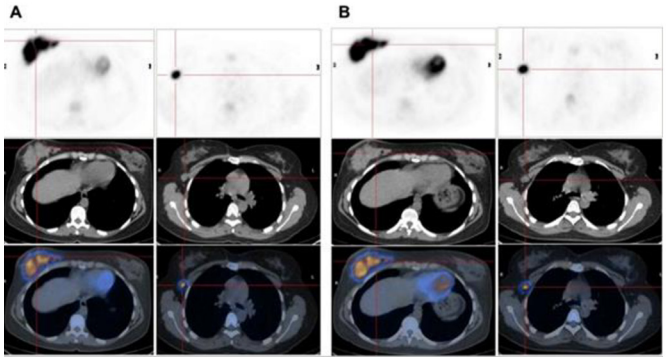

Figure 2. Female with right ductal invasive luminal B breast cancer. (a) Axial slices of baseline 18F-FDG PET/TC showing a significant increase of metabolism in the mammary lesion and lymph node. (b) Axial slices of interim 18F-FDG PET/TC with absence of a metabolic response. After neoadjuvant treatment, surgery revealed no histopathologic response (pT3N2). Interim PET/CT correctly classified the absence of a final response.

To establish the degree of metabolic response in solid tumors, different criteria have been described by the European Organization for Research and Treatment of Cancer (EORTC), the PET Response Criteria in Solid Tumors (PERCIST), and the latest revision of the RECIST 1.1, which includes the findings regarding FDG PET/CT [30-32]. Several meta-analyses have examined the effectiveness of FDG PET/CT regarding the prediction of favorable responses (pCR or nearly pCR) to NAT in patients with breast cancer that can be grouped into: (i) studies with PET/CT [33-36] (Table 1) and (ii) direct comparisons of PET/CT and MRI [29, 37-39] (Table 2). In the context of early response assessment, the pooled sensitivity and specificity values range from 80% to 88% and from 70% to 87%, respectively [34-36, 39].

Table 1. Effectiveness of FDG PET/CT regarding the prediction of favorable responses: studies with PET/CT

Author (year)	Aim and methodology	Techniques/ studies/patients	Results	Limitations and conclusions
Tian et al. (2017) [36]	Diagnostic performance in the prediction of NAC response for histopathologically confirmed BC patients (updated November 2016). Inclusion criteria: N>20 patients. Metabolic change measured by ΔSUVmax. Studies based on per patient statistics. Gold standard: Pathological response involving both breast tissues and/or lymph nodes. Subgroup analysis: (i) pCR vs responders as the reference standard, (ii) blinded vs non blinded PET results, (iii) prospective/retrospective design, (iv) timing performing PET (=2 or ≠ 2 cycles), (v) cut-off value (Δ SUVmax ≥ or <50%).	PET/CT 22 studies (17 studies after 1-2 cycles) 1119 patients	Pooled Se of 0.81 (0.76.0–0.86) and pooled Sp of 0.79 (0.72–0.85). DOR: 17.35 (10.98–27.42). Subgroup analysis: Only the specificity in study design comparison was significantly different (P < 0.05). No differences in Se or Sp regarding the timing of performing PET, with a pooled Se of 0.85 and a pooled Sp of 79%.	18F-FDG PET/CT has a moderate Acc in predicting the pathological response during the early process of NAC in BC patients. Limitations: Clinical heterogeneity derived from different pathological types or different stages. Nonhomogeneous subgroups of patients could cause publication bias, verification bias and selection bias.
Mghanga et al. (2013) [35]	Diagnostic performance in the prediction of early NAC response (updated June 2012). Inclusion criteria: Availability to calculate statistics for diagnostic parameters. Meeting abstracts excluded. Gold standard: Pathological response in breast tissues. Subgroup analysis: 1 cycle vs 2 cycles	PET or PET/CT 15 studies -1 cycle (7 studies) -2 cycles (8 studies) 745 patients	Pooled Se of 0.80 (95% CI, 0.75-0.84) and pooled Sp of 78.8 (95% CI, 74.1-83.0). PPV 79.8%, NPV 79.5% Early response analysis: - 1 cycle: pooled Se of 0.73 (0.67-0.79) and pooled Sp of 0.85 (0.79-0.90). DOR: 17.64 (7.87-39.54). - 2 cycles: pooled Se of 0.76 (0.69-0.82) and Sp of 0.84% (0.78-0.89). DOR: 19.70 (9.36-41.46).	FDG-PET has moderately high Se and Sp in the early detection of responders versus nonresponders. Limitations: Studies with very small samples influence the statistical power of the individual study and lead to imprecise and inconclusive results. Different tumor types and various stages of BC and evaluation responses at different courses of chemotherapy may result in clinical heterogeneity among studies. Some studies included both patients with partial response and cPC as responders.
Cheng et al. (2012) [33]	Diagnostic performance in the prediction of early NAC response (updated April 2011). Inclusion criteria: Studies based on per patient statistics. n> 10 patients, FDG, availability to calculate statistics for diagnostic parameters. Meetings abstracts excluded. Gold standard: Pathological response in breast tissues. Subgroup analysis: - PET vs PET/TC - Pathological response rate (≥50% vs <50%) - QUADAS quality score (≥12 vs <12), - Publication year (≥vs <2009).	PET or PET/TC 17 studies -PET/TC (10 studies) -PET (7 studies) 781 patients	Pooled Se of 0.84 (95 CI, 0.79–0.87) and pooled Sp of 0.71 (95% CI 0.66–0.75). Subgroup analysis: - PET/CT: Pooled Se of 0.84 (95% CI 0.79–0.89) and pooled Sp of 0.66 (95% CI 0.59–0.72). DOR: 17.63 (95% CI 7.43–41.82). AUC: 0.89. - PET: Pooled Se 0.82 (95% CI 0.74–0.89) and pooled Sp 0.79 of (95% CI 0.719–0.849). DOR: 13.64 (95% CI 7.43–25.03). AUC: 0.88. For methodological quality, DORs of subgroups with QUADAS quality scores ≥12 were higher than subgroups with quality scores <12. For publication year, statistically significant differences were found in Sp, DOR and AUC between different subgroups.	PET/CT and PET have reasonable Se in evaluating response to NAC in breast cancer; however, the Sp is relatively low. The combination of other imaging methods with FDG PET/CT or PET is recommended. Limitations: Inclusion of various types, stages of breast cancer and treatments that determine clinical heterogeneity among the studies. Subgroup analysis can partially eliminate the effect of heterogeneity, selection bias, publication bias, verification bias, or work-up difference between studies.
Wang et al. (2012) [34]	Diagnostic performance in the prediction of NAC response in breast and lymph nodes and exploration of its optimal regimen for clinical use (updated February 2011). Inclusion criteria: N> 10 patients. Gold standard: primary BC or regional lymph nodes pathological response. Subgroup analysis: - Studies with pCR as a reference standard. - Lymph node response. - Early (1-2 cycles) vs late (>3 cycles) valuation response. - Different cutoff SUV criteria (40-45% vs 65-65% vs>65%).	PET 19 studies -Breast (16 studies, 786 patients). -Lymph nodes (4 studies, 150 patients). -pCR as reference (5 studies, 920 patients).	Breast: Pooled Se of 0.84 (95% CI, 0.78–0.88) and pooled Sp of 66 (95% CI, 0.62–0.70). DOR: 11.90 (95% CI, 6.33–22.36). Subgroup analysis: -pCR as reference: Se 0.84 (95% CI, 0.71–0.93), Sp 0.64 (95% CI, 0.57–0.71); DOR: 8.59 (95% CI, 3.39–21.75). -Regional lymph nodes: Pooled Se of 0.92 (95% CI, 0.83–0.97) and pooled NPV of 0.88 (0.76–0.95). -Early: Pooled Se 0.88, Sp 0.70 Acc 0.76. Late: Pooled Se 0.81, Sp 0.61 and Acc 0.65. Early PET was significantly better than later (Acc 76% vs. 65%, p = 0.001). -The best correlation with pathology was by employing a reduction rate cut-off value of SUV between 55 and 65%.	PET is useful to predict NAC response in BC. However, the relatively low Sp and PPV still call for caution. Limitations: Heterogeneity among several studies could not be eliminated by subgroup analysis. Spectrum biases were possible since 14 articles included mainly certain pathological types of BC.

Acc: accuracy, AUC: area under the curve, BC: breast cancer, CI: confidence interval, CT: computed tomography, DOR: diagnostic odds ratio, FDG: 18F-fluorodeoxyglucose, N: number, NAC: neoadjuvant chemotherapy, NPV: negative predictive value, pCR: pathological complete response, PET: positron emission tomography, PPV: positive predictive value, QUADAS: quality assessment of diagnostic accuracy studies, Se: sensitivity, Sp: specificity, SUV: standardized uptake value, Δ: variation.

Table 2. Effectiveness of FDG PET/CT regarding the prediction of favorable responses: direct comparisons of PET/CT and MRI

Author (year)	Aim and methodology	Techniques/studies/ patients	Results	Limitations and conclusions
Li et al. (2018) [29]	Comparison of PET/CT (or PET) and MRI after preoperative NAC (updated February 2017). Other inclusion criteria: - N ≥ 10 patients. - Postoperative pathologic result as the gold standard (pCR vs non pCR). Defined Se as the capacity to detect residual tumors after NAC. - Availability to calculate statistical diagnostic parameters. - Semiquantitative parameters of MRI and PET images.	Total: 13 studies (11 full text; 2 abstracts). -PET/CT or PET (618 p). -MRI (575 p).	- PET: pooled Se of 0.77 (95% CI: 0.58-0.90) and pooled Sp of 0.78 (95% CI: 0.63-0.88). AUC 0.84 (*). - MRI: pooled Se of 0.88 (95% CI: 0.78-0.94) and pooled Sp of 0.69 (95% CI: 0.51-0.83). AUC 0.88 (*).	MRI had a higher sensitivity and PET/CT a higher specificity in predicting the pathologic response (*). Methodological heterogeneity: (i) differences in thresholds for response definitions of MRI and PET technologies; (ii) N cycles to monitor early response, differences between definitions of pCR used (sometimes not reported).
Chen et al. (2017) [39]	Comparison of PET/CT (or PET) and MRI accuracy after preoperative NAC (updated July 2016). Other inclusion criteria: - Postoperative pathologic result as the gold standard. - Study outcome as pCR or near-pCR to NAC. - Direct comparative design or randomized controlled trial. Subgroup analyses: (i) different cut-off values; (ii) different MRI protocols; (iii) different time points for early evaluation before and after 3 cycles.	Total: 11 studies -PET/CT or PET (527 p). -MRI (527 p).	- PET/CT: pooled Se of 0.87 (95% CI: 0.71–0.95) and pooled Sp of 0.85 (95% CI: 0.70–0.93). - MRI: pooled Se of 0.79 (95% CI: 0.68–0.87) and pooled Sp of 0.82 (95% CI: 0.72–0.89)(†). Subgroup analyses: PET/CT is superior to MRI in assessing response at times between 1–3 cycles of NAC, with a pooled Sp of 0.94 (95% CI: 0.78-0.98) vs. 0.83 (95% CI: 0.81- 0.87), p= 0.015), respectively, but not after 3 cycles of NAC.	- The diagnostic performance of MRI is similar to that of PET/CT for the assessment of BC response to NAC. However, PET/CT has more Se than MRI and more Sp if the second imaging scan is performed before 3 cycles of NAC. Limitations: Small sample sizes of comparative studies. Differences between definition of pCR and BC phenotype.
Sheikhbahaei et al. (2016) [38]	To establish the diagnostic performances of both MRI and FDG-PET/CT imaging for predicting residual disease after NAC. Other inclusion criteria: - Postoperative pathologic result as the gold standard. - Study outcome as pCR vs no pCR to NAC. - Direct comparative design or randomized controlled trial. Subgroup analyses: (a) Intra-NAC assessment (2-4 cycles). (b) Post-NAC assessment.	Total: 10 studies (8 full text, 2 abstracts) -PET/CT or PET (535 p). -MRI (492 p). (a) Intra-NAC: 3 studies (256 p) (b) Post-NAC: 7 studies (PET:279 p; MRI: 236p).	-PET/CT: Pooled Se of 0.71 (95% CI: 0.52–0.85) and pooled Sp of 0.77 (95% CI: 0.58–0.89). The pooled Se was significantly higher in studies that used combined FDG-PET/CT imaging vs FDG-PET imaging alone (*). -MRI: pooled Se of 0.88 (95% CI: 0.76–0.95) and pooled Sp of 0.55 (95% CI: 0.41–0.68) (*). Subgroup analyses (*): -Intra-NAC pooled Se/pooled Sp (CI 95%): PET/CT: 0.69 (0.25–0.93)/0.91 (0.86–0.95)/ MRI: 0.42 (0.20–0.68)/0.89 (0.66–0.97) -Post-NAC pooled Se/pooled Sp (95% CI): PET/CT: 0.88 (0.73–0.95)/0.71(0.42–0.89) MRI: 0.63 (0.51–0.74)/0.88 (0.71–0.96).	FDG-PET/CT imaging outperformed MRI in intra-NAC assessment, whereas the overall performance of MRI was higher after completion of NAC before surgery. Limitations: -Lack of consensus of response criteria. -Unable to perform subtype analysis. The timing of imaging for NAC response assessment exerts a major influence on the estimates of diagnostic accuracy.

AUC: area under the curve, BC: breast cancer, CI: confidence interval, CT: computed tomography, FDG: 18F-fluorodeoxyglucose, MRI: magnetic resonance imaging, N: number, NAC: neoadjuvant chemotherapy, pCR: pathological complete response, PET: positron emission tomography, Se: sensitivity, Sp: specificity, SUV: standardized uptake value, Δ: variation. (*): Diagnostic statistical parameters defined for the prediction of residual tumor. (†): Diagnostic statistical parameters defined for the prediction of response.

Compared with MRI, FDG PET/CT has greater sensitivity regarding the prediction of response to NAT between one to three cycles [29, 38, 39], although no agreement exists regarding specificity. Li et al. [29] described the superiority of MRI over FDG PET/CT, whereas Chen et al. [39] showed better results with FDG PET/CT.

A day-to-day collaboration between diagnostic and treatment teams will accelerate the application of novel improvements in imaging to decision making.

QUESTION 3: final assessment

Which are the preferred imaging methods for end of treatment evaluation?

Consensus recommendation
The preferred method for response estimation and pCR prediction is MRI with functional assessment, unless PET-CT have been used as the initial and intermediate evaluation method.

Literature review and clinical interpretation
The accurate determination of residual disease after the end of NAT might help decide the type of surgery as well as make long-term predictions of prognosis.

Mammography (MG) has shown a 74-79% predictive rate of pCR [40]. Its limitations include the lack of correlation between the presence of microcalcifications with viable disease and the adequate definition of margins, especially regarding spiculated lesions, which can be improved with the use of a digital tomosynthesis that decreases the masking effects of healthy tissue and improves its definition capacity. Ultrasound (US) improves prediction with respect to MG, especially at the nodal level, leading to an 80% prediction of pCR combined with MG [41]. However, the experience with both techniques is limited.

Most of the literature agrees that DCE-MRI is an accurate breast imaging technique for evaluating the extent of residual disease after NAT, although it depends on the lesion type and the response pattern to treatment. However, Vriens et al. [42] emphasized that US can be at least as effective as MRI in assessing the size of the residual lesion. A debate exists with the utility of MG after NAT in cases where microcalcifications (MCCs) persist after NAT. We know that post-NAT residual MCCs are not correlated with the presence of viable tumor; however, the absence of enhancement on MRI is significantly correlated with pCR. Feliciano et al. [43] suggested that complete excision of all indeterminate or malignant-appearing MCCs remains standard practice and is a substantial limitation to the use of NAT for downstaging of patients to breast conservative therapy.

Croshaw et al. [44] retrospectively reviewed the accuracy, positive predictive value (PPV), and negative predictive value (NPV) of DCE-MRI techniques in determining pCR in patients with breast cancer after NAT. They noted that all modalities had a high PPV greater than 75% for identifying the presence of residual disease; however, the NPV was low (less than 50%) and was the highest for MRI NPV, at 44%. These authors also performed a meta-analysis that evaluated 6 independent studies assessing the diagnostic properties of breast imaging techniques in determining pCR, obtaining similar data. All methods adequately predicted the presence of residual disease, with MRI being the most accurate (MRI PPV of 93%). However, the only method with a high NPV was MRI, at 65%. Thus, MRI is the best imaging technique for predicting residual disease and pCR [44].

With respect to the functional assessment of DCE-MRI and DWI-MRI, the most prominent meta-analyses agree that DCE-RM has a high specificity to detect pCR, and DWI-MRI has a high sensitivity to detect pCR [22, 45-47].

Published in July 2017, a meta-analysis conducted by Gu et al. [47] examined 60 studies that compared the DWI-MRI and DCE-MRI sequences to detect tumor responses to NAT (54 DCE-MRI studies and 8 DWI-MR studies). They observed that DCE-MRI had a pooled sensitivity of 64%, a pooled specificity of 93% and an AUC (area under the curve) of 88%, whereas DWI-MRI had a pooled sensitivity of 92%, a pooled specificity of 85% and an AUC of 94%; thus, the metanalysis indicated that DCE-MRI has a high specificity and DWI-MRI a high sensitivity.

Recent studies have analyzed the accuracy of quantitative kinetic parameters of DCE-MRI to detect pCR after NAT. According to different studies, different parameters correlate with pCR, such as the reduction of more than 98% in volume enhancement, the combination of a reduction of > 27% in SI (signal axis peak) and of >78% in area, and functional tumor volume with different threshold values, with the highest correlation with pCR [48-50].

Regarding response assessment by FDG PET/CT, the available metanalyses show mixed results. Tian et al. [36] found similar diagnostic performances for response assessment regarding the timing of PET. Nevertheless, Chen et al. [39] found higher specificity in the prediction of early NAT response patients, whereas Sheikhbahaei et al. [38] found higher specificity after completion of NAT (Figure 3 PET).

Compared with PET/CT, MRI is superior to PET/TC for assessing final response [38, 39].

QUESTION 4: evaluation of the Regional Lymph Node response

Which are the best methods for the assessment of regional lymph node response?

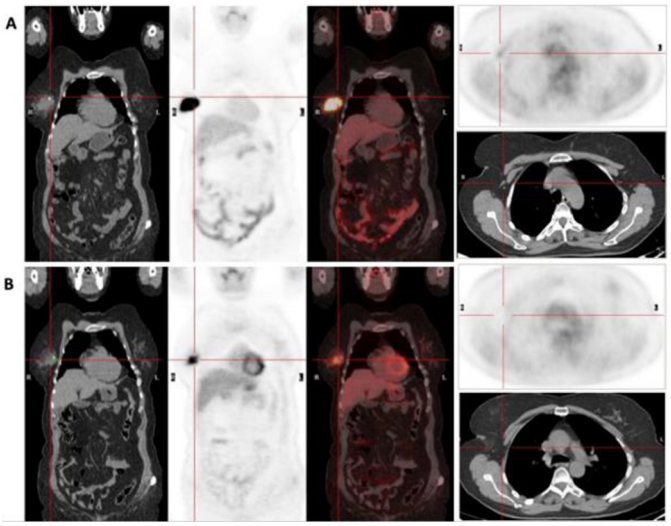

Figure 3. Female with right luminal B HER-2 (-) breast cancer with lymph node involvement. (a) Coronal and axial slices of baseline 18F-FDG PET/TC showing a significant increase of metabolism in the breast lesion and a mild increase in the axillary lymph node. (b) Coronal and axial slices of post-treatment PET/CT showing reduced metabolism and size of the mammary lesion, with no pathologic metabolism in the axillary lymph node. Surgery revealed a partial response in the breast lesion and a complete response in the lymph node. PET/CT correctly classified both responses.

Consensus recommendation
SLNB is the standard method for regional lymph node evaluation after NAT, but ecography and MRI should be complementary, especially in initially N+ patients.

Literature review and clinical interpretation
Numerous studies have examined the capacity of imaging for the evaluation of the axillary response to avoid unnecessary invasive procedures among patients exhibiting a possible pCR in the axilla after NAT [51]. In a study of 47 patients with lymph node metastases who underwent MRI before and after NAT, the sensitivity and specificity values regarding the detection of residual disease were 86% and 89%, respectively [52]. A prospective multimodal analysis showed that US had the highest sensitivity for the detection of axillary metastases after NAT (70%), followed by FDG PET/CT (63%) and MRI (61%) [53]. Alvarado et al. [54] showed that normalization of the morphological characteristics of lymph nodes after NAT strongly correlated with pCR.

Several classifications to characterize lymph nodes as pathological have sought to improve the performance and interobserver variability of US and have been used both at diagnosis and after NAT to provide a valid response assessment. The classification by Bedi considers the morphological characteristics of the lymph nodes and divides them into six categories. Type 5 (focal hypoechoic cortical lobulation) and type 6 (totally hypoechoic node with no hilum) are the most robust. Amonkar's study reduced these categories to four but added the quantification of cortical thickening, where the categories UN4 (cortex with uniform cortical thickness >=2.3 mm) and UN5 (enlarged node with no fatty hilum) are the most accurate [55, 56] (Figure 4 and Table 3)

Table 3. Amonkar classification. Nodal score according to morphological features on ultrasound

Nodal score	Morphological features
UN2 (normal)	Uniform cortex <2.3 mm and centrally placed fatty hilum
UN3 (indeterminate)	>=2.3 mm cortex with uniform cortical thickness
UN4 (suspicious)	Localized bulge of cortex >2.3 mm, eccentric displacement of fatty hilum, small vessels entering cortex of node (color flow ultrasound)
UN5 (replaced)	Enlarged node with no fatty hilum

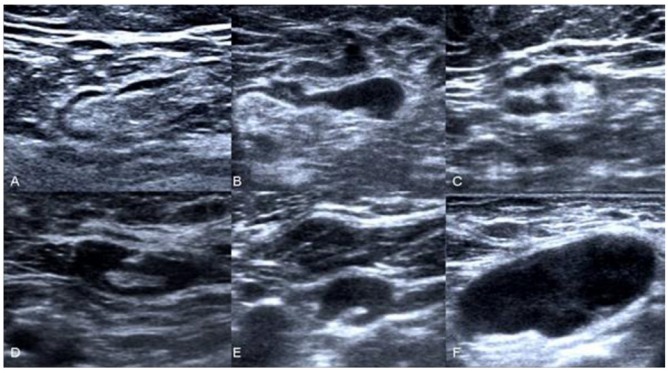

Figure 4. Morphologic ultrasound characteristics of metastatic axillary lymph nodes. a-Normal lymph node with fatty hilum and thin cortex. b, c, d, e, f -Metastatic lymph nodes. b and c- Asymmetrical focal cortical lobulation. d- Diffuse nodal cortical enlargement, FNB ultrasound-guided procedure. e- Diffuse nodal cortical enlargement. f- Absence of fatty hilum. This is the morphologic finding with higher PPV.

Despite these advances, axillary imaging cannot yet replace surgical staging. As of today, the fundamental role of current axillary imaging in NAT is related to decreasing the false-negative rate (FNR) of post-NAT sentinel lymph node biopsy (SLNB), which varies between 9.6% and 14.2% according to the clinical trials SNFNAT, ACOSOG Z1071, and SENTINA [57-59]. In this scenario, US helps to classify patients for axillary surgery, decreasing the FNR to 9.8% [58]. RECIST 1.1 only included the short-axis lymph node measurement to evaluate the response in post-NAT assessments. However, morphological findings, such as cortical thickening and the loss of fatty hilum, are equally important for determining residual disease; in fact, the loss of fatty hilum has the highest PPV [60].

Following the description of a correlation between the pathological lymph node percutaneously biopsied at diagnosis and the resection during the sentinel lymph node dissection (SLND) after NAT [61], several US-guided studies were launched, in which a clip is placed on the pathological node at diagnosis. After NAT, the marker is surgically removed during the SLND. In a subgroup of patients of the ACOSOG Z1071 trial whose nodes were marked, an FNR of 6.8% was obtained when the marked node was one of those included in the sentinel surgery, while the FNR was 19% if it was not present [62]. The

National Comprehensive Cancer Network (NCCN) guidelines commented on this technique to decrease the FNR during post-neoadjuvant SLND [63]. Different image-guided techniques are used to ensure the inclusion of the marked node during SLND. The final objective of these techniques is the realization of a target axillary dissection (TAD) that involves sufficient axillary surgery and fewer lesions than lymphadenectomy [64].

New data are indicating the possibility of avoiding axillary surgery altogether after conversion to N- after NAT; imaging is going to have a crucial role in the decision-making process [65-67].

The most-used techniques that have proven their efficacy thus far in clinical trials make use of radioactive iodine seeds both prior to NAT and after treatment. Alternatively, presurgical localization with wires of the lymph nodes previously marked with biopsy markers is employed to obtain a post-neoadjuvant FNR between 1.4%-7% [64, 68-70]. Other methods of labeling in the breast appear promising for the post-neoadjuvant armpit, including carbon suspension, magnetic seeds, radar reflectors, and radiofrequency identification devices [71].

Information is limited regarding the role of FDG PET/CT in the evaluation of lymph node response, although a high pooled sensitivity and a negative predictive value (92% for both parameters) have been documented [34].

As an important note of caution, detection of a pathological lymph node by imaging should not lead to automatic fine-needle aspiration or core needle biopsy and reflex complete axillary dissection if positive. Several authors have shown that, in the adjuvant setting, Z011 criteria should prevail [72, 73]. Several studies are ongoing in the neoadjuvant setting.

FACTORS AFECTING THE DIAGNOSTIC PERFORMANCE OF IMAGING

Several factors can modify the diagnostic accuracy of MRI and FDG PET/CT and must therefore be considered:

Non-standardized definition of pCR
There is not a standard pCR definition among studies. Moreover, the criteria used to evaluate histological responses are diverse, as some studies include partial responders in the favorable response group. Thus, MRI and FDG PET/CT accuracies depend on whether the definition includes the presence or not of carcinoma in situ (CIS) and the axillary lymph node status.

Non-standardized definition for imaging complete response (iCR)
There is no standard iCR definition. For radiological imaging, the most commonly used scales are the Response Evaluation Criteria in Solid Tumors (RECIST) and the Union for International Cancer Control (UICC); however, none of these scales include DWI-MRI evaluation. The classic definition of iCR is the lack of enhancement, but no agreement exists about whether it should be measured in the early or late phase or if it should be compared with background parenchyma enhancement (BPE).

Regarding FDG PET/CT, highly variable reduction rates are used across different studies, which mostly range between 40% and 80%. Some even use the EORTC criteria or 5-point scale evaluations [34-36, 39]. However, the best correlation with pathology is yielded by employing a reduction rate cut-off value of SUV between 55% and 65% [34].

Lesion type on pre-NAT and tumor response pattern
For MRI, when a single nodular enhancement or a minimal BPE is present, there is a greater correlation with pCR. The pathological correlation is reduced when CIS, lobular carcinoma, coil artifact or a fragmented response pattern is present (Figure 5).

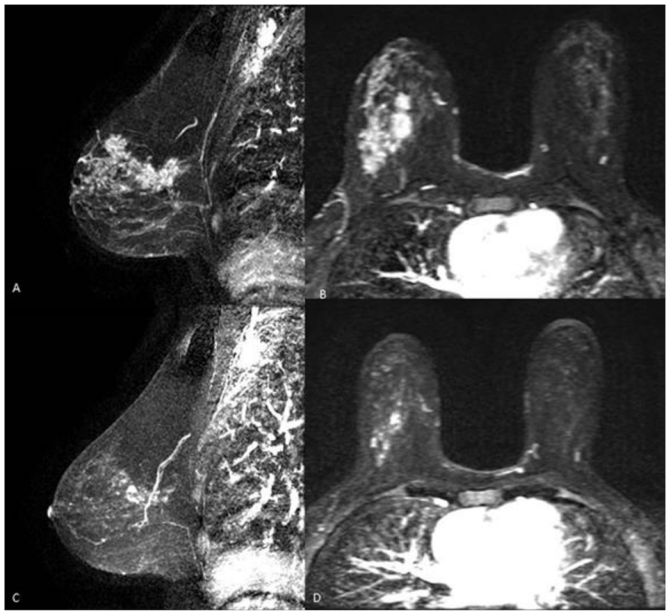

Figure 5. Images of a 50-year-old woman with newly diagnosed clinical stage IIB (T3N0M0) right breast cancer. Histologic results were grade 3, ER-positive, PR-positive, and HER2-nonamplifed invasive ductal carcinoma. (a, b) Maximum intensity projection from gadolinium-enhanced breast MR imaging performed prior to neoadjuvant chemotherapy demonstrates an irregular right breast mass with heterogeneous enhancement and non-mass enhancement extending anterior to the mass. (c, d) After neoadjuvant chemotherapy shows a crumbling response with little change in major diameter. Final pathologic findings after mastectomy showed good correlations.

The degree of tumor avidity by FDG PET/CT is a fundamental factor in the assessment of metabolic response. Tumors with low avidity for FDG at baseline are more difficult to distinguish from normal tissue; therefore, the assessment of the metabolic response is less precise [29].

Molecular subtypes and tumor grade
Molecular subtypes are determinants of the degree of metabolic activity as well as of the treatment regimen and overall response.

With respect to MRI, Houssami conducted a meta-analysis of 30 studies and 12,000 patients and found a statistical and independent association between tumor subtype and pCR, with a higher pCR rate associated with HER2 and TN subtypes [74] (Figure 6).

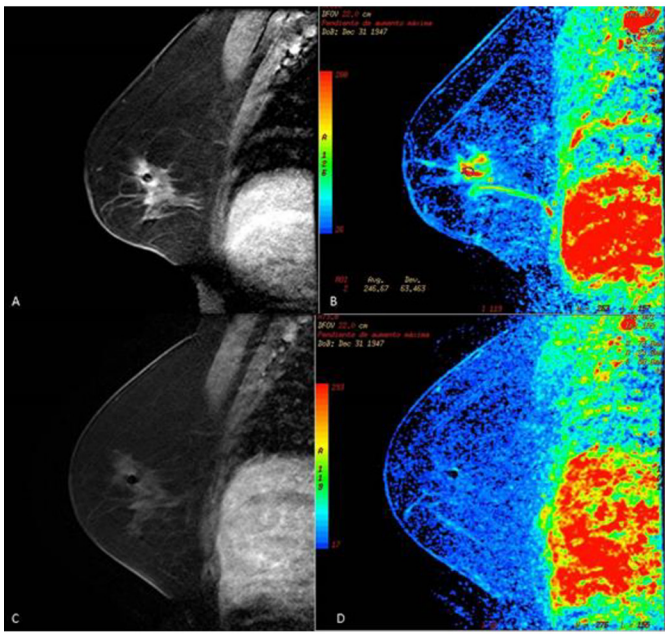

Figure 6. Images of a 40-year-old woman with newly diagnosed clinical stage IIA (T2N0M0) right breast cancer. Histologic results were grade 2, ER-negative, PR-negative, and HER2-negative invasive ductal carcinoma. (a, b) Sagittal MRI T1WI post-gadolinium demonstrates an irregular mass of 2.5 cm with heterogeneous enhancement. Kinetic map shows increased perfusion. (c, d) NAC MRI demonstrates a complete response, which correlates with the histologic findings after lumpectomy.

In 2007, Namura et al. [75] investigated the predictive values (PVs) based on molecular subtype. This author considered 6 molecular subgroups differentiated by HR+ subtypes into strongly positive and moderately positive and separated the PPV based on a strict definition of pCR and the definition of pCR that allows CIS. The global PPV was 46% (without differentiation by subtype), but pure TN and HER2 subtypes had higher values, with a TN PPV of 58% and a HER2 PPV of 56%. When the pCR definition included CIS, the PPVs increased to 73% and 92%, respectively (Figure 7). With respect to DWI-MRI, some authors have concluded that a combined model of CE-MRI and MRI-DWI increases diagnostic accuracy [76, 77].

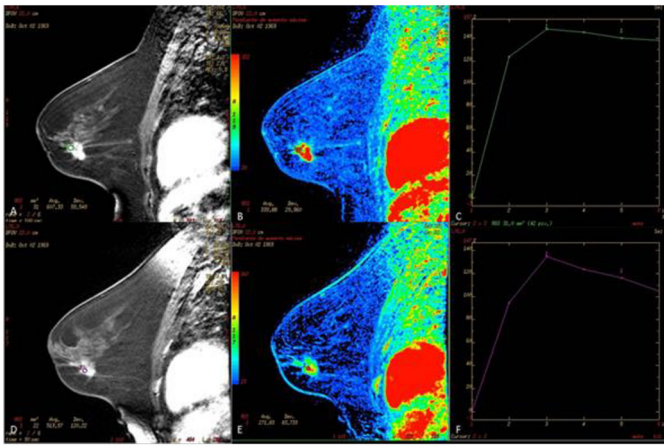

Figure 7. Images of a 65-year-old woman with newly diagnosed clinical stage IIA (T2N0M0) left breast cancer. Histologic results were grade 2, ER-positive, PR-positive, and HER2-negative invasive ductal carcinoma. (a, b, c) Sagittal MRI T1WI post-gadolinium demonstrates an irregular mass of 2 cm with heterogeneous enhancement. Kinetic map shows increased perfusion. (d, e, f) NAC MRI demonstrates no change in the tumor size but some kinetic changes, with a decrease in perfusion. Histological findings after lumpectomy showed a 2-cm malignant nodule.

Using FDG PET/CT, significant differences were observed in the diagnostic accuracy depending on the molecular subtype [27, 29]. Thus, although no specific response criteria exist for each histological subtype, the most aggressive subtypes, e.g., TN and HER2+, present with higher metabolic activity than HR+HER2- [78]. In TN tumors, the Δ% SUVmax can be used to predict the pCR, with an accuracy greater than 75% [77]. Regarding the HER2+ subtype, PET/CT presents accuracies between 56% and 90% regarding the prediction of pCR. The results are less promising for the subtypes HR+HER2-, which have lower FDG uptake [79].

Type of NAT and timing of interim imaging during NAT

Using MRI, taxanes in weekly schedules and anti-angiogenic drugs can produce false-positive iCR due to their anti-vascular effects [80].

For FDG PET/CT, when regimens with dense or intense doses are administered, the variations observed for Δ% SUVmax are generally greater than conventional dose regimens, thereby improving the accuracy of PET/CT to assess responses. Thus, a cut-off point of Δ% SUVmax to distinguish between responders and non-responders should be established specifically for each molecular subtype and chemotherapy regimen used [27].

The assessment of early response after one to two cycles is more accurate than after three or more cycles (76% vs 65%, p=0.01) [34]. Thus, the most favorable studies are those that determine the early response using PET/CT after two cycles of NAT, evaluating the residual activity in the breast and/or axilla [34-36].

All the previous conditions explain the variability of the results observed using the different techniques, justifying their difficulty for integration into clinical practice [27, 29].

In relation to determining the superiority of one technique versus another, large-scale, head-to-head, well-designed trials using a common methodology will be necessary to compare the predictive values of PET/CT and MRI that consider factors such as the definition of pCR and the phenotype of the breast cancer [39].

CONCLUSIONS

Both MRI and FDG PET/CT are effective tools, with moderate-to-high sensitivity and specificity in the evaluation of the response to NAT. Based on the reported evidence, both MRI and FDG PET/CT can be recommended for NAT prediction of response, with moderate recommendation (IIa) and moderate quality of evidence (level B).

Although the lack of standardization of pCR and iCR definitions significantly limits the interpretation of the different studies, PET/CT seems to perform better in the early prediction after one to three cycles and MRI in later evaluations (after three or more cycles of NAT). However, the diagnostic techniques should be selected based on the availability and diagnostic experience of the imaging team at each site in close collaboration with the clinical team. There is still significant room to improve the predictive and prognostic capacities of the imaging methods associated with NAT. Clinical decisions associated with earlier response detection can lead to better results and less toxicity. Thus, the incorporation of functional imaging methods, primarily DWI, DCE perfusion, spectroscopy within the field of MRI and FDG PET/CT, provides more precise images of complex tumor biological processes, such as increased proliferation, inhibited apoptosis, metabolic patterns, neoangiogenesis, immune response, and their changes during NAT. These methods will allow us to understand treatment responses beyond strictly volumetric criteria. Given the continuous and rapid evolution of imaging approaches, breast cancer biology and treatments, and the different equipment available at each center, a close collaboration between clinical and imaging experts would optimize the application of advances in each field into clinical care.

LIST OF ABBREVIATIONS

ADC, apparent diffusion coefficient
AUC, area under the curve
BC, breast cancer
BPE, background parenchyma enhancement
CI, confidence interval
CIS, carcinoma in situ
CT, computed tomography
DCE, dynamic contrast enhanced
DOR, diagnostic odds ratio
DWI, diffusion weighted imaging
EORTC, European Organization for Research and Treatment of Cancer
ER, estrogen receptor
FDA, Food and Drug Administration
FDG, fluorodeoxyglucose
FNR, false-negative rate
HRs, hormone receptors
iCR, imaging complete response
LAR, luminal with androgen expression
MCCs, microcalcifications
MG, mammography
MRI, magnetic resonance imaging
N, number
NAC, neoadjuvant chemotherapy
NAT, neoadjuvant treatment of breast cancer
NCCN, The National Comprehensive Cancer Network
NPV, negative predictive value
pCR, pathologic complete response
PET, positron emission tomography
PPV, positive predictive value
PR, progesterone receptor
QUADAS, quality assessment of diagnostic accuracy studies
RCB, residual cancer burden
RECIST, response evaluation criteria in solid tumors
Se, sensitivity
SEMNIM, Spanish Society of Nuclear and Molecular Imaging
SEOM, Spanish Society of Medical Oncology
SERAM, Spanish Society of Medical Radiology
SI, signal axis peak
SLNB, sentinel lymph node biopsy
SLND, sentinel lymph. node dissection
Sp, specificity
SUV, standardized uptake value
TAD, target axillary dissection
TN, triple-negative
UICC, Union for International Cancer Control

DECLARATIONS

Ethics approval and consent to participate
The manuscript does not contain any studies with human participants or animals performed by any of the authors. Informed consent/ethical approval is not required for this type of project.

Consent for publication
Not applicable

Availability of data and material
Not applicable

Competing interests
Ana María García-Vicente, Myriam Montes Fernandez, Ricardo Sánchez-Escribano, Ana Paula Caresia-Aróztegui, Sara Jiménez Arranz, Pedro Sánchez Rovira, Aurora Crespo de la Jara, and Francisco Javier de Castro-García report no conflict of interest. José E. Alés-Martínez has received speaker honorarium from Roche, MSD, and BMS; he has received consulting honorarium from Tesaro, Pfizer; and he has received travel grants from MSD, BMS, Roche.

Funding
This consensus was equally funded by the Spanish Society of Medical Oncology (SEOM), Spanish Society of Medical Radiology (SERAM), and Spanish Society of Nuclear Medicine and Molecular Imaging (SEMNIM).

Authors' contributions
All authors have contributed equally to the content of this manuscript

Acknowledgements
The authors would like to thank Fernando Rico-Villademoros (COCIENTE SL, Madrid, Spain) for editorial assistance in the preparation of this manuscript. His participation has been funded by the participating scientific societies.

REFERENCES

1. Early Breast Cancer Trialists' Collaborative Group (EBCTCG) (2018) Long-term outcomes for neoadjuvant versus adjuvant chemotherapy in early breast cancer: meta-analysis of individual patient data from ten randomised trials. Lancet Oncol 19:27-39.
2. Cortazar P, Zhang L, Untch M, Mehta K, Costantino JP, Wolmark N, Bonnefoi H, Cameron D, Gianni L, Valagussa P, Swain SM, Prowell T, Loibl S, Wickerham DL, Bogaerts J, Baselga J, Perou C, Blumenthal G, Blohmer J, Mamounas EP, Bergh J, Semiglazov V, Justice R, Eidtmann H, Paik S, Piccart M, Sridhara R, Fasching PA, Slaets L, Tang S, Gerber B, Geyer CE, Jr., Pazdur R, Ditsch N, Rastogi P, Eiermann W, von Minckwitz G (2014) Pathological complete response and long-term clinical benefit in breast cancer: the CTNeoBC pooled analysis. Lancet 384:164-172.
3. Symmans WF, Wei C, Gould R, Yu X, Zhang Y, Liu M, Walls A, Bousamra A, Ramineni M, Sinn B, Hunt K, Buchholz TA, Valero V, Buzdar AU, Yang W, Brewster AM, Moulder S, Pusztai L, Hatzis C, Hortobagyi GN (2017) Long-term prognostic risk after neoadjuvant chemotherapy associated with residual cancer burden and breast cancer subtype. J Clin Oncol 35:1049-1060.
4. von Minckwitz G, Untch M, Blohmer JU, Costa SD, Eidtmann H, Fasching PA, Gerber B, Eiermann W, Hilfrich J, Huober J, Jackisch C, Kaufmann M, Konecny GE, Denkert C, Nekljudova V, Mehta K, Loibl S (2012) Definition and impact of pathologic complete response on prognosis after neoadjuvant chemotherapy in various intrinsic breast cancer subtypes. J Clin Oncol 30:1796-1804.
5. Masuda H, Baggerly KA, Wang Y, Zhang Y, Gonzalez-Angulo AM, Meric-Bernstam F, Valero V, Lehmann BD, Pietenpol JA, Hortobagyi GN, Symmans WF, Ueno NT (2013) Differential response to neoadjuvant chemotherapy among 7 triple-negative breast cancer molecular subtypes. Clin Cancer Res 19:5533-5540.
6. Gianni L, Pienkowski T, Im YH, Roman L, Tseng LM, Liu MC, Lluch A, Staroslawska E, de la Haba-Rodriguez J, Im SA, Pedrini JL, Poirier B, Morandi P, Semiglazov V, Srimuninnimit V, Bianchi G, Szado T, Ratnayake J, Ross G, Valagussa P (2012) Efficacy and safety of neoadjuvant pertuzumab and trastuzumab in women with locally advanced, inflammatory, or early HER2-positive breast cancer (NeoSphere): a randomised multicentre, open-label, phase 2 trial. Lancet Oncol 13:25-32.
7. Broglio KR, Quintana M, Foster M, Olinger M, McGlothlin A, Berry SM, Boileau JF, Brezden-Masley C, Chia S, Dent S, Gelmon K, Paterson A, Rayson D, Berry DA (2016) Association of pathologic complete response to neoadjuvant therapy in HER2-positive breast cancer with long-term outcomes: a meta-analysis. JAMA Oncol 2:751-760.
8. Bae MS, Shin SU, Ryu HS, Han W, Im SA, Park IA, Noh DY, Moon WK (2016) Pretreatment MR imaging features of triple-negative breast cancer: association with response to neoadjuvant chemotherapy and recurrence-free survival. Radiology 281:392-400.
9. Mukhtar RA, Yau C, Rosen M, Tandon VJ, Hylton N, Esserman LJ (2013) Clinically meaningful tumor reduction rates vary by prechemotherapy MRI phenotype and tumor subtype in the I-SPY 1 TRIAL (CALGB 150007/150012; ACRIN 6657). Ann Surg Oncol 20:3823-3830.
10. Braman NM, Etesami M, Prasanna P, Dubchuk C, Gilmore H, Tiwari P, Plecha D, Madabhushi A (2017) Intratumoral and peritumoral radiomics for the pretreatment prediction of pathological complete response to neoadjuvant chemotherapy based on breast DCE-MRI. Breast Cancer Res 19:57.
11. Goldhirsch A, Wood WC, Coates AS, Gelber RD, Thürlimann B, Senn HJ (2011) Strategies for subtypes--dealing with the diversity of breast cancer: highlights of the St. Gallen international expert consensus on the primary therapy of early breast cancer 2011. Ann Oncol 22:1736-1747.
12. Buchbender C, Kuemmel S, Hoffmann O, Stahl AR, Kimmig R, Otterbach F, Ladd S, Koeninger A, Forsting M, Bockisch A, Antoch G, Heusner TA (2012) FDG-PET/CT for the early prediction of histopathological complete response to neoadjuvant chemotherapy in breast cancer patients: initial results. Acta Radiol 53:628-636.
13. Ueda S, Saeki T, Shigekawa T, Omata J, Moriya T, Yamamoto J, Osaki A, Fujiuchi N, Misumi M, Takeuchi H, Sakurai T, Tsuda H, Tamura K, Ishida J, Abe Y, Imabayashi E, Kuji I, Matsuda H (2012) 18F-fluorodeoxyglucose positron emission tomography optimizes neoadjuvant chemotherapy for primary breast cancer to achieve pathological complete response. Int J Clin Oncol 17:276-282.
14. Vicente AMG, Mora MÁC, Martín AAL, Mdel MMS, Calatayud FR, López OVG, Aunión RE, Ageitos AG, Castrejón AS (2014) Glycolytic activity with 18F-FDG PET/CT predicts final neoadjuvant chemotherapy response in breast cancer. Tumour Biol 35:11613-11620.
15. Kolesnikov-Gauthier H, Vanlemmens L, Baranzelli MC, Vennin P, Servent V, Fournier C, Carpentier P, Bonneterre J (2012) Predictive value of neoadjuvant chemotherapy failure in breast cancer using FDG-PET after the first course. Breast Cancer Res Treat 131:517-525.
16. Koolen BB, Pengel KE, Wesseling J, Vogel WV, Vrancken Peeters MJ, Vincent AD, Gilhuijs KG, Rodenhuis S, Rutgers EJ, Olmos RAV (2013) FDG PET/CT during neoadjuvant chemotherapy may predict response in ER-positive/HER2-negative and triple negative, but not in HER2-positive breast cancer. Breast 22:691-697.
17. García-Esquinas MAG, García JA, García-Sáenz JA, Furió-Bacete V, Ferrer MEF, Candil AO, Martín MNC, Delgado JLC (2014) Predictive value of PET-CT for pathological response in stages II and III breast cancer patients following neoadjuvant chemotherapy with docetaxel. Rev Esp Med Nucl Imagen Mol 33:14-21.
18. Soussan M, Orlhac F, Boubaya M, Zelek L, Ziol M, Eder V, Buvat I (2014) Relationship between tumor heterogeneity measured on FDG-PET/CT and pathological prognostic factors in invasive breast cancer. PLoS One 9:e94017.
19. Ha S, Park S, Bang JI, Kim EK, Lee HY (2017) Metabolic radiomics for pretreatment 18F-FDG PET/CT to characterize locally advanced breast cancer: histopathologic characteristics, response to neoadjuvant chemotherapy, and prognosis. Sci Rep 7:1556.

20. Moscoso A, Ruibal Á, Domínguez-Prado I, Fernández-Ferreiro A, Herranz M, Albaina L, Argibay S, Silva-Rodríguez J, Pardo-Montero J, Aguiar P (2018) Texture analysis of high-resolution dedicated breast (18) F-FDG PET images correlates with immunohistochemical factors and subtype of breast cancer. Eur J Nucl Med Mol Imaging 45:196-206.
21. Molina-García D, García-Vicente AM, Pérez-Beteta J, Amo-Salas M, Martínez-González A, Tello-Galán MJ, Soriano-Castrejón Á, Pérez-García VM (2018) Intratumoral heterogeneity in (18)F-FDG PET/CT by textural analysis in breast cancer as a predictive and prognostic subrogate. Ann Nucl Med 32:379-388.
22. Marinovich ML, Houssami N, Macaskill P, Sardanelli F, Irwig L, Mamounas EP, von Minckwitz G, Brennan ME, Ciatto S (2013) Meta-analysis of magnetic resonance imaging in detecting residual breast cancer after neoadjuvant therapy. J Natl Cancer Inst 105:321-333.
23. Marinovich ML, Sardanelli F, Ciatto S, Mamounas E, Brennan M, Macaskill P, Irwig L, von Minckwitz G, Houssami N (2012) Early prediction of pathologic response to neoadjuvant therapy in breast cancer: systematic review of the accuracy of MRI. Breast 21:669-677.
24. Hylton NM, Blume JD, Bernreuter WK, Pisano ED, Rosen MA, Morris EA, Weatherall PT, Lehman CD, Newstead GM, Polin S, Marques HS, Esserman LJ, Schnall MD (2012) Locally advanced breast cancer: MR imaging for prediction of response to neoadjuvant chemotherapy--results from ACRIN 6657/I-SPY TRIAL. Radiology 263:663-672.
25. Padhani AR, Hayes C, Assersohn L, Powles T, Makris A, Suckling J, Leach MO, Husband JE (2006) Prediction of clinicopathologic response of breast cancer to primary chemotherapy at contrast-enhanced MR imaging: initial clinical results. Radiology 239:361-374.
26. Fatayer H, Sharma N, Manuel D, Kim B, Keding A, Perren T, Velikova G, Lansdown M, Shaaban AM, Dall B (2016) Serial MRI scans help in assessing early response to neoadjuvant chemotherapy and tailoring breast cancer treatment. Eur J Surg Oncol 42:965-972.
27. Groheux D (2018) Role of fludeoxyglucose in breast cancer: treatment response. PET Clin 13:395-414.
28. Baizán AN, Puig DR, Segura JP (2018) Evolution of quantification methods in oncologic (18)F-FDG PET studies. Rev Esp Med Nucl Imagen Mol 37:203-204.
29. Li H, Yao L, Jin P, Hu L, Li X, Guo T, Yang K (2018) MRI and PET/CT for evaluation of the pathological response to neoadjuvant chemotherapy in breast cancer: a systematic review and meta-analysis. Breast 40:106-115.
30. Therasse P, Arbuck SG, Eisenhauer EA, Wanders J, Kaplan RS, Rubinstein L, Verweij J, Van Glabbeke M, van Oosterom AT, Christian MC, Gwyther SG (2000) New guidelines to evaluate the response to treatment in solid tumors. J Natl Cancer Inst 92:205-216.
31. Eisenhauer EA, Therasse P, Bogaerts J, Schwartz LH, Sargent D, Ford R, Dancey J, Arbuck S, Gwyther S, Mooney M, Rubinstein L, Shankar L, Dodd L, Kaplan R, Lacombe D, Verweij J (2009) New response evaluation criteria in solid tumours: revised RECIST guideline (version 1.1). Eur J Cancer 45:228-247.
32. Wahl RL, Jacene H, Kasamon Y, Lodge MA (2009) From RECIST to PERCIST: evolving considerations for PET response criteria in solid tumors. J Nucl Med 50:122S-150S.
33. Cheng X, Li Y, Liu B, Xu Z, Bao L, Wang J (2012) 18F-FDG PET/CT and PET for evaluation of pathological response to neoadjuvant chemotherapy in breast cancer: a meta-analysis. Acta Radiol 53:615-627.
34. Wang Y, Zhang C, Liu J, Huang G (2012) Is 18F-FDG PET accurate to predict neoadjuvant therapy response in breast cancer? A meta-analysis. Breast Cancer Res Treat 131:357-369.
35. Mghanga FP, Lan X, Bakari KH, Li C, Zhang Y (2013) Fluorine-18 fluorodeoxyglucose positron emission tomography-computed tomography in monitoring the response of breast cancer to neoadjuvant chemotherapy: a meta-analysis. Clin Breast Cancer 13:271-279.
36. Tian F, Shen G, Deng Y, Diao W, Jia Z (2017) The accuracy of (18)F-FDG PET/CT in predicting the pathological response to neoadjuvant chemotherapy in patients with breast cancer: a meta-analysis and systematic review. Eur Radiol 27:4786-4796.
37. Liu Q, Wang C, Li P, Liu J, Huang G, Song S (2016) The role of (18)F-FDG PET/CT and MRI in assessing pathological complete response to neoadjuvant chemotherapy in patients with breast cancer: a systematic review and meta-analysis. Biomed Res Int 2016:3746232.
38. Sheikhbahaei S, Trahan TJ, Xiao J, Taghipour M, Mena E, Connolly RM, Subramaniam RM (2016) FDG-PET/CT and MRI for evaluation of pathologic response to neoadjuvant chemotherapy in patients with breast cancer: a meta-analysis of diagnostic accuracy studies. Oncologist 21:931-939.
39. Chen L, Yang Q, Bao J, Liu D, Huang X, Wang J (2017) Direct comparison of PET/CT and MRI to predict the pathological response to neoadjuvant chemotherapy in breast cancer: a meta-analysis. Sci Rep 7:8479.
40. Huber S, Wagner M, Zuna I, Medl M, Czembirek H, Delorme S (2000) Locally advanced breast carcinoma: evaluation of mammography in the prediction of residual disease after induction chemotherapy. Anticancer Res 20:553-558.
41. Peintinger F, Kuerer HM, Anderson K, Boughey JC, Meric-Bernstam F, Singletary SE, Hunt KK, Whitman GJ, Stephens T, Buzdar AU, Green MC, Symmans WF (2006) Accuracy of the combination of mammography and sonography in predicting tumor response in breast cancer patients after neoadjuvant chemotherapy. Ann Surg Oncol 13:1443-1449.
42. Vriens BE, de Vries B, Lobbes MB, van Gastel SM, van den Berkmortel FW, Smilde TJ, van Warmerdam LJ, de Boer M, van Spronsen DJ, Smidt ML, Peer PG, Aarts MJ, Tjan-Heijnen VC (2016) Ultrasound is at least as good as magnetic resonance imaging in predicting tumour size post-neoadjuvant chemotherapy in breast cancer. Eur J Cancer 52:67-76.
43. Feliciano Y, Mamtani A, Morrow M, Stempel MM, Patil S, Jochelson MS (2017) Do calcifications seen on mammography after neoadjuvant chemotherapy for breast cancer always need to be excised? Ann Surg Oncol 24:1492-1498.
44. Croshaw R, Shapiro-Wright H, Svensson E, Erb K, Julian T (2011) Accuracy of clinical examination, digital mammogram, ultrasound, and MRI in determining postneoadjuvant pathologic tumor response in operable breast cancer patients. Ann Surg Oncol 18:3160-3163.
45. Wu LM, Hu JN, Gu HY, Hua J, Chen J, Xu JR (2012) Can diffusion-weighted MR imaging and contrast-enhanced MR imaging precisely evaluate and predict pathological response to neoadjuvant chemotherapy in patients with breast cancer? Breast Cancer Res Treat 135:17-28.
46. Lobbes MB, Prevos R, Smidt M, Tjan-Heijnen VC, van Goethem M, Schipper R, Beets-Tan RG, Wildberger JE (2013) The role of magnetic resonance imaging in assessing residual disease and pathologic complete response in breast cancer patients receiving neoadjuvant chemotherapy: a systematic review. Insights Imaging 4:163-175.
47. Gu YL, Pan SM, Ren J, Yang ZX, Jiang GQ (2017) Role of magnetic resonance imaging in detection of pathologic complete remission in breast cancer patients treated with neoadjuvant chemotherapy: a meta-analysis. Clin Breast Cancer 17:245-255.
48. Li YL, Zhang XP, Li J, Cao K, Cui Y, Li XT, Sun YS (2015) MRI in diagnosis of pathological complete response in breast cancer patients after neoadjuvant chemotherapy. Eur J Radiol 84:242-249.
49. Li W, Arasu V, Newitt DC, Jones EF, Wilmes L, Gibbs J, Kornak J, Joe BN, Esserman LJ, Hylton NM (2016) Effect of MR imaging contrast thresholds on prediction of neoadjuvant chemotherapy response in breast cancer subtypes: a subgroup analysis of the ACRIN 6657/I-SPY 1 TRIAL. Tomography 2:378-387.
50. Lo WC, Li W, Jones EF, Newitt DC, Kornak J, Wilmes LJ, Esserman LJ, Hylton NM (2016) Effect of imaging parameter thresholds on MRI prediction of neoadjuvant chemotherapy response in breast cancer subtypes. PLoS One 11:e0142047.
51. Chang JM, Leung JWT, Moy L, Ha SM, Moon WK (2020) Axillary nodal evaluation in breast cancer: state of the art. Radiology 295:500-515.
52. Javid S, Segara D, Lotfi P, Raza S, Golshan M (2010) Can breast MRI predict axillary lymph node metastasis in women undergoing neoadjuvant chemotherapy. Ann Surg Oncol 17:1841-1846.

53. Hieken TJ, Boughey JC, Jones KN, Shah SS, Glazebrook KN (2013) Imaging response and residual metastatic axillary lymph node disease after neoadjuvant chemotherapy for primary breast cancer. Ann Surg Oncol 20:3199-3204.
54. Alvarado R, Yi M, Le-Petross H, Gilcrease M, Mittendorf EA, Bedrosian I, Hwang RF, Caudle AS, Babiera GV, Akins JS, Kuerer HM, Hunt KK (2012) The role for sentinel lymph node dissection after neoadjuvant chemotherapy in patients who present with node-positive breast cancer. Ann Surg Oncol 19:3177-3184.
55. Bedi DG, Krishnamurthy R, Krishnamurthy S, Edeiken BS, Le-Petross H, Fornage BD, Bassett RL, Jr., Hunt KK (2008) Cortical morphologic features of axillary lymph nodes as a predictor of metastasis in breast cancer: *in vitro* sonographic study. AJR Am J Roentgenol 191:646-652.
56. Amonkar SJ, Oates E, McLean L, Nicholson S (2013) Pre-operative staging of the axilla in primary breast cancer. By redefining the abnormal appearing node can we reduce investigations without affecting overall treatment? Breast 22:1114-1118.
57. Kuehn T, Bauerfeind I, Fehm T, Fleige B, Hausschild M, Helms G, Lebeau A, Liedtke C, von Minckwitz G, Nekljudova V, Schmatloch S, Schrenk P, Staebler A, Untch M (2013) Sentinel-lymph-node biopsy in patients with breast cancer before and after neoadjuvant chemotherapy (SENTINA): a prospective, multicentre cohort study. Lancet Oncol 14:609-618.
58. Boileau JF, Poirier B, Basik M, Holloway CM, Gaboury L, Sideris L, Meterissian S, Arnaout A, Brackstone M, McCready DR, Karp SE, Trop I, Lisbona A, Wright FC, Younan RJ, Provencher L, Patocskai E, Omeroglu A, Robidoux A (2015) Sentinel node biopsy after neoadjuvant chemotherapy in biopsy-proven node-positive breast cancer: the SN FNAC study. J Clin Oncol 33:258-264.
59. Boughey JC, Ballman KV, Hunt KK, McCall LM, Mittendorf EA, Ahrendt GM, Wilke LG, Le-Petross HT (2015) Axillary ultrasound after neoadjuvant chemotherapy and its impact on sentinel lymph node surgery: results from the American college of surgeons oncology group Z1071 trial (Alliance). J Clin Oncol 33:3386-3393.
60. Garcia-Ortega MJ, Benito MA, Vahamonde EF, Torres PR, Velasco AB, Paredes MM (2011) Pretreatment axillary ultrasonography and core biopsy in patients with suspected breast cancer: diagnostic accuracy and impact on management. Eur J Radiol 79:64-72.
61. Nathanson SD, Burke M, Slater R, Kapke A (2007) Preoperative identification of the sentinel lymph node in breast cancer. Ann Surg Oncol 14:3102-3110.
62. Boughey JC, Ballman KV, Le-Petross HT, McCall LM, Mittendorf EA, Ahrendt GM, Wilke LG, Taback B, Feliberti EC, Hunt KK (2016) Identification and resection of clipped node decreases the false-negative rate of sentinel lymph node surgery in patients presenting with node-positive breast cancer (T0-T4, N1-N2) who receive neoadjuvant chemotherapy: results from ACOSOG Z1071 (Alliance). Ann Surg 263:802-807.
63. National Comprehensive Cancer Network (2015) National Comprehensive Cancer Network (NCCN) guidelines version 2.2015. Preoperative systemic therapy breast and axillary evaluation. National Comprehensive Cancer Network, Philadelphia.
64. Caudle AS, Yang WT, Krishnamurthy S, Mittendorf EA, Black DM, Gilcrease MZ, Bedrosian I, Hobbs BP, DeSnyder SM, Hwang RF, Adrada BE, Shaitelman SF, Chavez-MacGregor M, Smith BD, Candelaria RP, Babiera GV, Dogan BE, Santiago L, Hunt KK, Kuerer HM (2016) Improved axillary evaluation following neoadjuvant therapy for patients with node-positive breast cancer using selective evaluation of clipped nodes: implementation of targeted axillary dissection. J Clin Oncol 34:1072-1078.
65. Chehade HEH, Headon H, El Tokhy O, Heeney J, Kasem A, Mokbel K (2016) Is sentinel lymph node biopsy a viable alternative to complete axillary dissection following neoadjuvant chemotherapy in women with node-positive breast cancer at diagnosis? An updated meta-analysis involving 3,398 patients. Am J Surg 212:969-981.
66. Choi HJ, Kim I, Alsharif E, Park S, Kim JM, Ryu JM, Nam SJ, Kim SW, Yu J, Lee SK, Lee JE (2018) Use of sentinel lymph node biopsy after neoadjuvant chemotherapy in patients with axillary node-positive breast cancer in diagnosis. J Breast Cancer 21:433-441.
67. Cerbelli B, Botticelli A, Pisano A, Campagna D, De Vincentiis L, Pernazza A, Frusone F, Scavina P, Monti M, Fortunato L, Costarelli L, d'Amati G (2019) Breast cancer subtypes affect the nodal response after neoadjuvant chemotherapy in locally advanced breast cancer: are we ready to endorse axillary conservation? Breast J 25:273-277.
68. Donker M, Straver ME, Wesseling J, Loo CE, Schot M, Drukker CA, van Tinteren H, Sonke GS, Rutgers EJ, Peeters MJV (2015) Marking axillary lymph nodes with radioactive iodine seeds for axillary staging after neoadjuvant systemic treatment in breast cancer patients: the MARI procedure. Ann Surg 261:378-382.
69. van Nijnatten TJA, Simons JM, Smidt ML, van der Pol CC, van Diest PJ, Jager A, van Klaveren D, Kam BLR, Lobbes MBI, de Boer M, Verhoef K, Koppert LB, Luiten EJT (2017) A novel less-invasive approach for axillary staging after neoadjuvant chemotherapy in patients with axillary node-positive breast cancer by combining radioactive iodine seed localization in the axilla with the sentinel node procedure (RISAS): a dutch prospective multicenter validation study. Clin Breast Cancer 17:399-402.
70. Hartmann S, Reimer T, Gerber B, Stubert J, Stengel B, Stachs A (2018) Wire localization of clip-marked axillary lymph nodes in breast cancer patients treated with primary systemic therapy. Eur J Surg Oncol 44:1307-1311.
71. Woods RW, Camp MS, Durr NJ, Harvey SC (2019) A review of options for localization of axillary lymph nodes in the treatment of invasive breast cancer. Acad Radiol 26:805-819.
72. Alés-Martínez JE, Tur R, Parra J, Martín MdR, Filipovich E, Segovia B, Ceballos J, Revestido R, De Castro J, De Grado C (2018) Relationship of axillary total tumoral load (TTL) in early breast cancer and local and distant clinical outcomes. J Clin Oncol 36:e12574-e12574.
73. Jozsa F, Ahmed M, Baker R, Douek M (2019) Is sentinel node biopsy necessary in the radiologically negative axilla in breast cancer? Breast Cancer Res Treat 177:1-4.
74. Houssami N, Macaskill P, von Minckwitz G, Marinovich ML, Mamounas E (2012) Meta-analysis of the association of breast cancer subtype and pathologic complete response to neoadjuvant chemotherapy. Eur J Cancer 48:3342-3354.
75. Namura M, Tsunoda H, Yagata H, Hayashi N, Yoshida A, Morishita E, Takei J, Suzuki K, Yamauchi H (2018) Discrepancies between pathological tumor responses and estimations of complete response by magnetic resonance imaging after neoadjuvant chemotherapy differ by breast cancer subtype. Clin Breast Cancer 18:128-134.
76. Bufi E, Belli P, Di Matteo M, Terribile D, Franceschini G, Nardone L, Petrone G, Bonomo L (2014) Effect of breast cancer phenotype on diagnostic performance of MRI in the prediction to response to neoadjuvant treatment. Eur J Radiol 83:1631-1638.
77. Santamaría G, Bargalló X, Fernández PL, Farrús B, Caparrós X, Velasco M (2017) Neoadjuvant systemic therapy in breast cancer: association of contrast-enhanced MR imaging findings, diffusion-weighted imaging findings, and tumor subtype with tumor response. Radiology 283:663-672.
78. Groheux D, Mankoff D, Espié M, Hindié E (2016) ^{18}F-FDG PET/CT in the early prediction of pathological response in aggressive subtypes of breast cancer: review of the literature and recommendations for use in clinical trials. Eur J Nucl Med Mol Imaging 43:983-993.
79. Cheng J, Wang Y, Mo M, Bao X, Zhang Y, Liu G, Zhang J, Geng D (2015) 18F-fluorodeoxyglucose (FDG) PET/CT after two cycles of neoadjuvant therapy may predict response in HER2-negative, but not in HER2-positive breast cancer. Oncotarget 6:29388-29395.
80. Schrading S, Kuhl CK (2015) Breast cancer: influence of taxanes on response assessment with dynamic contrast-enhanced MR imaging. Radiology 277:687-696.