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

Ana María García-Vicente[^1], Myriam Montes Fernandez[^1,2], Ricardo Sánchez-Escribano[^3,4], Ana Paula Caresia-Ároztegui[^3,4], Sara Jiménez Arranz[^3,4], Pedro Sánchez Roviira[^3,4], Aurora Crespo de la Jarra[^3,4], Francisco Javier de Castro-García[^3,4] and José E. Álés-Martínez[^3,4]

^1Nuclear Medicine Department, University General Hospital of Ciudad Real, Calle Obispo Rafael Torija, s/n, 13005 Ciudad Real, Spain, E-mail: angarvice@yahoo.es / ^2Radiology Department, Hospital Clínico San Carlos, Calle Professor Martin Lagos, s/n, 28040 Madrid, Spain, E-mail: myriammfag@hotmail.com / ^3Medical Oncology Department, Hospital Clínico Universitario de Valladolid (HCUV), Avenida Ramón y Cajal, s/n, 47003 Valladolid, Spain, E-mail: rsescrbano@seom.org / ^4Radiology Department, Hospital Nuestra Señora de Sonsoles, Avenida Juan Carlos I, s/n, 05004 Ávila, Spain, E-mail: fjcastro@comunidade.org / ^5Radiology Department, Hospital Universitario 12 de Octubre, Avenida de Córdoba, s/n, 28041 Madrid, Spain, E-mail: saragimenezarranz@gmail.com / ^6Medical Oncology Department, Hospital Universitario de Jaén, Avenida del Ejército Español, 10, 23007 Jaén, Spain, E-mail: onccpsr@yahoo.es / ^7Medical Oncology Department, University Hospital of San Juan, Carretera N-352, s/n, 03550 San Juan de Alicante, Spain, E-mail: crespo_aur@gva.es / ^8Radiology Department, Hospital Nuestra Señora de Sonsoles, Avenida Juan Carlos I, s/n, 05004 Ávila, Spain, E-mail: fcojav7769@gmail.com / ^9Medical Oncology Department, Hospital Nuestra Señora de Sonsoles, Avenida Juan Carlos I, s/n, 05004 Ávila, Spain, E-mail: jaleasm@seom.org / ^10Spanish Society of Nuclear Medicine and Molecular Imaging (SEMIMI) / ^11Spanish Society of Medical Radiology (SERAM) / ^12Spanish Society of Medical Oncology (SEOM)

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 (SEMIMI) 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.

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/ypN0) 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. Álés-Martínez, MD
Medical Oncology Department, Hospital Nuestra Señora de Sonsoles, Avda Juan Carlos I, S/n. 05004 Ávila, Spain / Tel: +34619287189 / Email: jaleasm@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 chemotheraphy, 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 oncoprotein 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 (SEMMIM) 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 position 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 late evaluations after two 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 metabollic 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<sub>max</sub> (value of the voxel with the highest SUV), although others have used SUV<sub>mean</sub> (mean of SUV voxels within the tumor volume), SUV<sub>ROI</sub> (mean of SUV voxels in a spherical region of interest of 1 cm<sup>3</sup> near the SUV<sub>max</sub> 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<sub>max</sub>), 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.
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

<table>
<thead>
<tr>
<th>Author (year)</th>
<th>Aim and methodology</th>
<th>Techniques/studies/patients</th>
<th>Results</th>
<th>Limitations and conclusions</th>
</tr>
</thead>
<tbody>
<tr>
<td>Tian et al. (2017) [38]</td>
<td>Diagnostic performance in the prediction of NAC response for histopathologically confirmed BC patients (updated November 2016). Inclusion criteria: N&gt;20 patients. Metabolic change measured by SUVmax. Studies based on per patient statistics.</td>
<td>PET/CT: 22 studies (17 studies after 1-2 cycles) 1119 patients</td>
<td>Pooled Se of 0.81 (0.76-0.86) and pooled Sp of 0.79 (0.72-0.85). DOR: 17.35 (10.98-27.42).</td>
<td>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%.</td>
</tr>
</tbody>
</table>

| Mghanaga et al. (2013) [35] | Diagnostic performance in the prediction of early NAC response (updated June 2012). | 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% | FDG-PET has moderately high Se and Sp in the early detection of responders versus nonresponders. |

| Cheng et al. (2012) [33] | Diagnostic performance in the prediction of early NAC response (updated April 2011). Inclusion criteria: Studies based on patient statistics. N>10 patients, FDG availability to calculate statistics for diagnostic parameters. Meetings abstracts excluded. | 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). | PET/CT and PET have reasonable Se in evaluating response to NAT in breast cancer; however, the Sp is relatively low. The combination of other imaging methods with PET/CT or PET is recommended. |

| Wang et al. (2012) [34] | Diagnostic performance in the prediction of NAC response in breast and lymph nodes and evaluation of its optimal regimen for clinical use (updated February 2011). | PET: 19 studies Breast (15 studies, 789 patients). Lymph nodes (4 studies, 150 patients) | Breast: Pooled Se of 0.84 (95% CI, 0.79-0.88) and pooled Sp of 0.80 (95% CI, 0.75-0.85). DOR: 11.90 (5.73-25.03). | PET is useful in predicting NAT response in BC. However, the relatively low Sp and PPV still call for caution. |


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

<table>
<thead>
<tr>
<th>Author (year)</th>
<th>Aim and methodology</th>
<th>Techniques/studies/patients</th>
<th>Results</th>
<th>Limitations and conclusions</th>
</tr>
</thead>
<tbody>
<tr>
<td>Chen et al. (2017) [39]</td>
<td>Comparison of PET/CT (or PET) and MRI accuracy after neoadjuvant preoperative NAT (updated July 2016).</td>
<td>Total: 11 studies - PET/CT or PET (527 p). - MRI (527 p).</td>
<td>- 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 NAT, 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 NAT.</td>
<td>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).</td>
</tr>
<tr>
<td>Sheikhbahaei et al. (2016) [38]</td>
<td>To establish the diagnostic performances of both MRI and FDG-PET/CT imaging for predicting residual disease after NAT.</td>
<td>Total: 10 studies (8 full text, 2 abstracts) - PET/CT or PET (535 p). - MRI (492 p).</td>
<td>(a) Intra-NAC: 3 studies (256 p)</td>
<td>The diagnostic performance of MRI is similar to that of PET/CT for the assessment of BC response to NAT. However, PET/CT has more Se than MRI and more Sp if the second imaging scan is performed before 3 cycles of NAT. Limitations: Small sample sizes of comparative studies. Differences between definitions of pCR and BC phenotype.</td>
</tr>
<tr>
<td>Li et al. (2018) [29]</td>
<td>Comparison of PET/CT (or PET) and MRI after neoadjuvant preoperative NAT (updated February 2017).</td>
<td>Total: 13 studies (11 full text; 2 abstracts). - PET/CT or PET (618 p). - MRI (575 p).</td>
<td>- 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 (<em>). - MRI: pooled Se of 0.88 (95% CI: 0.79–0.94) and pooled Sp of 0.69 (95% CI: 0.51–0.83). AUC 0.88 (</em>).</td>
<td>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).</td>
</tr>
</tbody>
</table>

**AUC** area under the curve, **BC** breast cancer, **CI** confidence interval, **CT** computed tomography, **FDG** 18F-fluorodeoxyglucose, **MRI** magnetic resonance imaging, **N** number, **NAT** 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 predicions 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].
In a study of 47 patients (SLNB), which Uniform cortex <2.3 mm and centrally placed fatty hilum >=2.3 mm cortex with uniform cortical thickness varies between 9.6% and SLND) Enlarged node with no fatty hilum after NAT [61], several US-Node response QUESTION 4: evaluation of the Regional Lymph final response [38, 39]. Compared with PET/CT, MRI is superior to PET/CT for assessing completion of NAT (Figure 3 PET). Nevertheless, Chen et al. [39] found higher diagnostic performances for response assessment regarding the timing of PET. Regarding response assessment by FDG PET/CT, the available metaanalyses show mixed results. Tian et al. [36] found similar specificities 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?

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 breast lesion and a histologic lesion that is surgically removed 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 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 are 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

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

Table 3.

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

**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 breast lesion and a histologic lesion that is surgically removed 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 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).

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 SNF-NAT, 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 1.1.1 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 AFFECTING 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 and 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).
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].

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 (Ila) 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
SEMMIM, 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
REFERENCES


