Breast Cancer Subtype Influences the Accuracy of Predicting Pathologic Response by Imaging and Clinical Breast Exam After Neoadjuvant Chemotherapy

ASHLEY R. WALDREP1, ERIC J. AVERY2, FERRILL F. ROSE JR.1, MADHU V. MIDATHADA2, JONI A. TILFORD2, HANS-CHRISTIAN KOLBERG3 and MARK R. HUTCHINS2

1Union College Physician Assistant Program, Lincoln, NE, U.S.A.; 2Nebraska Hematology-Oncology, PC, Lincoln, NE, U.S.A.; 3Clinic for Gynecology and Obstetrics, Marienhospital Bottrop gGmbH, Bottrop, Germany

Abstract. Background/Aim: Clinical response evaluation after neoadjuvant chemotherapy (NACT) for breast cancer could include various imaging methods, as well as clinical breast exam (CBE). We assessed the accuracy of CBE and imaging to predict pathologic response after NACT administration according to breast cancer subtype. Patients and Methods: This retrospective cohort study included 84 patients with records of NACT and subsequent primary breast surgery from 2003-2013. Patients were divided into 4 breast cancer subtypes according to hormone receptor (HR) status and human epidermal growth factor receptor-2 (HER2) status. Negative predictive value (NPV), false-negative rate (FNR), false-positive rate (FPR) and positive predictive value (PPV) were calculated for CBE and imaging post-NACT and prior to breast cancer surgery. Results: NPV, FNR, FPR and PPV varied by breast cancer subtype and clinical response evaluation method. Imaging resulted in a higher NPV and a lower FNR than CBE among the entire cohort. There was a lower FPR with CBE. Clinical response evaluation by CBE was highly accurate for predicting pathologic residual disease in HR+ tumors (CBE PPV: 95.5% in HR+HER2−, 100.0% in HR+HER2+). In triple-negative breast cancer (TNBC), the imaging NPV was 100% and the imaging FNR was 0%. Conclusion: The use of imaging in HR+ tumors post-NACT may provide little to no additional value that is not already garnered by performance of a CBE. For TNBC, imaging may play a critical role in the prediction of pathologic complete response (pCR) post-NACT.

Neoadjuvant chemotherapy (NACT) for early breast cancer is equivalent to adjuvant chemotherapy in terms of disease-free survival and overall survival (1). The major difference between NACT and adjuvant chemotherapy is the ability to directly observe the treatment effect on the tumor. Clinical response can be assessed throughout the administration of NACT, while pathologic response is revealed upon breast surgery after completion of NACT. A pathologic complete response (pCR) after NACT, defined as no residual invasive breast cancer, translates into an improved long-term prognosis (2, 3).

The use of NACT also provides the opportunity to explore the potential value of response-guided therapy for early non-responders (4-6). An improvement in disease-free survival for hormone receptor (HR)-positive breast cancer using a response-guided NACT regimen has been reported previously (4). On the other end of the spectrum, it could hypothetically be beneficial to identify early responders, who may have reached a complete remission after a shorter duration of therapy than planned. For these patients, completing NACT to the planned number of cycles could be unnecessary and potentially detrimental in terms of toxicity and cost.

Routine clinical breast exam (CBE), as well as imaging modalities, such as mammography (MG), ultrasound (US) and magnetic resonance imaging (MRI), have been used to assess clinical response in an attempt to predict pathological response. A meta-analysis revealed that MRI most accurately predicts the amount of residual pathologic disease in non-subtyped early breast cancer patients when compared to CBE, MG and US (7). Yet, there have been conflicting reports on whether MRI typically underestimates or overestimates the amount of residual disease after NACT (8, 9). The routine use of MRI in clinical practice may at times be challenged by a patient’s pre-existing conditions, technical feasibility and
insurance coverage. A convenient and affordable method of clinical response evaluation is the CBE. It can be performed throughout the course of NACT administration for clinical response monitoring. Unfortunately, CBE less accurately predicts the amount of residual pathologic disease for non-subtyped early breast cancer in comparison to MRI (7).

Breast cancer subtyping by the status of HRs and human epidermal growth factor receptor-2 (HER2) form the basis of current systemic treatment planning and prediction of long-term prognosis. In early breast cancer patients who are candidates for chemotherapy, it is known that the rate of pCR after traditional NACT varies significantly by breast cancer subtype, with HR-negative tumors achieving higher rates of pCR than HR-positive tumors (10). For HER2-positive tumors, the use of HER2-targeted therapies significantly increases the rate of pCR (10).

As the treatment of early breast cancer continues to transform into a precision model of care, it is logical to investigate the accuracy of clinical response assessment methods according to breast cancer subtype, especially if those assessments could have an impact on decisions about further NACT administration. Recent studies have shown that the accuracy of imaging in clinical response evaluation varies by breast cancer subtype according to HR and HER2 status (11-19). Specifically, MRI more accurately predicts pathologic response in HR-negative subtypes rather than HR-positive subtypes (11, 16-19).

The purpose of this retrospective cohort analysis was to assess the ability of CBE and imaging to predict pCR/non-pCR after NACT according to breast cancer subtype. Although CBE is less accurate at predicting final pathology in non-subtyped studies, it is still a standard of care during the course of NACT. As the accuracy of MRI varies by breast cancer subtype after NACT, it is possible that the same could be true for CBE after NACT. By further evaluating the routine methods for clinical response evaluation according to inherent breast cancer subtype, we can improve the confidence with which personalized treatment recommendations are made during and after NACT.

**Patients and Methods**

This was a retrospective, exploratory cohort study of female breast cancer patients treated with NACT between 2003 and 2013 at Nebraska Hematology-Oncology, a single community oncology practice in Lincoln, NE, USA. This study was granted institutional review board exempt status by the Union College Human Subjects Review Board in accordance with 45CFR46.101(b) (4). Patients who had existing records of both neoadjuvant chemotherapy administration and final surgical pathology were included in the study. Eighty-four patients were identified. One patient was excluded from the study cohort upon further review due to harboring both HR-positive disease and TNBC within a single breast by separate needle biopsies, making it impossible to be able to assign a single subtype for the case (n=83). Another patient had bilateral synchronous breast cancer. Since we evaluated tumors and not cases, these two tumors were analyzed separately and this resulted in a final tumor number of 84.

Breast cancer subsets were designated as follows: HR-positive/HER2-negative (HR+HER2−), HR-positive/HER2-positive (HR+HER2+), HR-negative/HER2-positive (HR−/HER2+) and TNBC. Tumors with estrogen receptor (ER) and/or progesterone receptor (PgR) expression ≥10% by immunohistochemistry (IHC) or positive by reverse transcription polymerase chain reaction (RT-PCR) were considered HR+. For subset designation, the decision to use ≥10% by IHC as a definition of HR positivity is a method consistent with other similar studies (6, 7, 9). Tumors with HER2 expression of 3+ by IHC or 2+ by IHC with positive gene amplification by fluorescence in situ hybridization (FISH) were considered HER2+.

The NACT regimens received varied from patient to patient. They were generally standard protocols that could be administered in the post-operative adjuvant setting. All HER2-positive patients received HER2-targeting agents with their NACT regimen. Some HER2-positive patients received an anthracycline-based NACT regimen (9/24), while the remaining 15 received non-anthracycline regimens. Of the 38 HR+HER2− patients, 27 received an anthracycline-based NACT regimen, 8 received docetaxel/cyclophosphamide and 3 received docetaxel/carboplatin. Among the 22 TNBC patients, 12 received an anthracycline-based NACT regimen, 8 received platinum/taxane regimens and 1 patient received cyclophospha-mide/docetaxel.

Clinical stage at diagnosis was reported per the American Joint Committee on Cancer criteria and according to the largest primary tumor/lymph node measurements on clinical exam or imaging (20). pCR was defined as the absence of residual invasive cancer on pathologic evaluation of the resected breast specimen and all sampled regional lymph nodes (ypT0/Tis ypN0). Non-pathologic complete response (non-pCR) was then defined as presence of residual invasive cancer in breast or lymph node pathology.

Clinical complete response (cCR) or non-clinical complete response (non-cCR) by CBE and imaging were evaluated for prediction of pCR or non-pCR. By CBE, cCR was defined as the absence of residual mass, thickening or skin changes in the breast and axilla, which had been present at the baseline CBE. The CBE used in this analysis was typically conducted on day 1 of the final cycle of NACT. Five patients had a negative CBE at diagnosis and were not included in the CBE cCR analysis, decreasing that n to 79. By imaging, cCR was defined as the absence of residual breast cancer in the breast and axilla per radiology report. For each patient, the imaging result that was analyzed for this study was the one performed as close as possible prior to surgery. The most common imaging modality that was used was breast MRI (n=53). Other patients were evaluated by either MG, US, computed tomography (CT) or positron emission tomography (PET) scanning methods (n=13). Eighteen patients did not undergo imaging after NACT, decreasing the imaging cCR analysis n to 66.

Negative predictive value (NPV) was calculated as the number of patients clinically assessed as a cCR who resulted in a pCR divided by all patients clinically assessed as a cCR:

$$\text{NPV} = \frac{\text{cCR} \rightarrow \text{pCR}}{\text{all cCR}}$$

In the context of this study, a high NPV is desirable in order to predict that a cCR is truly a pCR. A limitation of the NPV is that a low prevalence of disease will increase the NPV, meaning that a high rate of pCR within the cohort will increase the NPV result.
False-negative rate (FNR) was calculated as the number of patients clinically assessed as a cCR who resulted in a non-pCR divided by all patients with a non-pCR:

$$\frac{cCR - non-pCR}{all \ non-pCR}$$

False-positive rate (FPR) was calculated as the number of patients clinically assessed as a non-cCR who resulted in a pCR divided by all patients with a pCR:

$$\frac{non-cCR - pCR}{all \ pCR}$$

A low FNR and FPR are desirable characteristics of a test used to predict final pathology.

Positive predictive value (PPV) was calculated as the number of patients clinically assessed as a non-cCR who resulted in a non-pCR divided by all patients clinically assessed as a non-cCR:

$$\frac{non-cCR - non-pCR}{all \ non-cCR}$$

A high PPV is desirable to accurately predict that a non-cCR is truly a non-pCR. The PPV is also a test influenced by disease prevalence, as a high prevalence of disease will increase the PPV, meaning that a low rate of pCR within the cohort will increase the PPV.

Rates of pCR were calculated for the cohort and each subset with 95% confidence intervals (CIs) by Wilson interval estimation for binomial proportions. The data analysis was generated using SAS University Edition software (http://www.sas.com/en_us/software/university-edition/download-software.html). Graphs were generated in Excel (http://www.microsoft.com/en-us/download/details.aspx?id=3) and exported in high-resolution with Daniel’s XL Toolbox add-in (21).

**Results**

The mean age of this retrospective cohort was 51.6 years. The median year of breast cancer diagnosis was 2010. HR+HER2- cancers comprised the greatest percentage of the cohort (45.2%). The majority of patients had ductal histology (81.0%) and the majority had grade III tumors (47.6%). Most patients had a clinical T2 primary tumor (47.6%) and half of all patients had clinically positive lymph node involvement at diagnosis. Refer to Table I for full details of the study cohort.

The rate of pCR for the entire cohort was 29.8% (95% CI=21.0%-40.3%), with the HR-HER2+ and TNBC subsets experiencing higher rates of pCR, respectively, at 53.9% (95% CI=29.1%-76.8%) and 40.9% (95% CI=23.3%-61.3%). Refer to Table II for each pCR rate.

Table III, Figure 1, and Figure 2 display the NPV, FNR, FPR and PPV of clinical response evaluation for the entire cohort and by each of the 4 breast cancer subtypes. Rates were calculated for both CBE and imaging clinical response evaluation methods. NPV, FNR, FPR and PPV varied by breast cancer subtype. Imaging did provide a higher NPV and a lower FNR than CBE among the entire cohort. However, there was a lower FPR with CBE. In the HR+HER2+ subtype, CBE was superior to imaging in all clinical response evaluation proportions except for FNR. Imaging had a NPV of 80% and a FNR of 25% in HR-HER2+ breast cancer. In TNBC, the imaging NPV was 100% and the imaging FNR was 0%.
Discussion

The patients analyzed in this investigation represented the typical high-risk cohort to be expected in the use of NACT. A recent retrospective review of the National Cancer Database revealed a pCR rate of 29.7% for breast cancer patients in the United States from 2010-2011, which matches our result of a pCR rate of 29.8% for the entire cohort (22). Our results revealed that the HR–HER2+ and TNBC subtypes achieved much higher rates of pCR than the HR+ subtypes.

To our knowledge, this is the first study of its kind to report the performance of CBE, as well as imaging, in prediction of pCR/non-pCR according to breast cancer subtype. In routine practice, it is likely that patients will be assessed by CBE throughout NACT and prior to being assessed by imaging. To illustrate the results of this retrospective study, if a patient was assessed by CBE as having a complete eradication of disease in the breast and axilla, that assessment was not very accurate (CBE NPV: 23.1% in HR+HER2−, 42.9% in HR+HER2+, 50% in HR–HER2+, 46.2% in TNBC). For patients who were assessed by imaging to have complete eradication of disease in the breast and axilla, the assessment was very accurate in certain subtypes (imaging NPV: 50.0% in HR+HER2−, 25.0% in HR+HER2+, 80.0% in HR–HER2+, 100.0% in TNBC). These results suggest the importance of using
imaging aside from CBE for prediction of pCR in TNBC and will be further discussed in following paragraphs.

HR−HER2+ breast cancer is known to have high rates of pCR, as we could reproduce in our study with a pCR rate of 53.9% (95% CI=29.1%-76.8%). The use of imaging, in this subtype, to predict pCR may also be important because the imaging NPV of 80.0% was clearly superior to the CBE NPV of only 50.0%.

The results for the HR+HER2− subtype also demonstrated an increased accuracy in the prediction of pCR by imaging vs. CBE (imaging NPV: 50.0% vs. CBE NPV: 23.1%) but, still, only half of the patients assessed as a cCR by imaging had a pCR.

The value of imaging for pCR prediction after NACT in HR+HER2+ breast cancer is questioned by our study results as imaging was less accurate at predicting pCR when compared to CBE (imaging NPV: 25% vs. CBE NPV: 42.9%). It should be noted, however, that CBE had a higher FNR than imaging (50.0% vs. 37.5% in HR+HER2+) and the sample size of the HR+HER2+ cohort was small (n=11). Our results lead to the conclusion that the HR+HER2+ subtype is the most inaccurately assessed subtype for prediction of pCR by imaging.

If a patient was assessed by CBE as having residual clinical disease in the breast or axilla, that assessment was highly accurate for predicting pathologic residual disease in HR+ tumors (CBE PPV: 95.5% in HR+HER2−, 100.0% in HR+HER2+ vs. imaging PPV: 95.0% in HR+HER2−, 71.4% in HR+HER2+). These results raise the question of the real benefit of additional imaging in HR+ tumors to assess for clinical residual disease. In addition, within our small cohort of HR+HER2+ disease, imaging had a poor FPR of 66.7% compared to CBE with a FPR of 0%.

In HR- tumors, the CBE PPV was nearly identical to the imaging PPV (CBE PPV: 50.0% in HR−HER2+, 75.0% in TNBC vs. imaging PPV: 50.0% in HR−HER2+, 78.6% in TNBC). However, imaging did have a higher FPR than CBE in the HR- tumors.

Table II. Rates of pathologic complete response.

<table>
<thead>
<tr>
<th></th>
<th>N</th>
<th>pCR% (95% CI)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Total</td>
<td>84</td>
<td>29.8 (21.0-40.3)</td>
</tr>
<tr>
<td>HR+HER2−</td>
<td>38</td>
<td>15.8 (7.4-30.4)</td>
</tr>
<tr>
<td>HR+HER2+</td>
<td>11</td>
<td>27.3 (9.8-56.6)</td>
</tr>
<tr>
<td>HR−HER2+</td>
<td>13</td>
<td>53.9 (29.1-76.8)</td>
</tr>
<tr>
<td>TNBC</td>
<td>22</td>
<td>40.9 (23.3-61.3)</td>
</tr>
</tbody>
</table>

Table III. NPV, FNR, PPV and FPR of CBE and imaging in predicting pCR/non-pCR by breast cancer subtype.

<table>
<thead>
<tr>
<th></th>
<th>CBE</th>
<th>Imaging</th>
</tr>
</thead>
<tbody>
<tr>
<td>Total</td>
<td>79</td>
<td>66</td>
</tr>
<tr>
<td>NPV (%)</td>
<td>16/41 (39.0)</td>
<td>12/19 (63.2)</td>
</tr>
<tr>
<td>FN (%)</td>
<td>25/58 (43.1)</td>
<td>7/45 (15.6)</td>
</tr>
<tr>
<td>PPV (%)</td>
<td>33/38 (86.8)</td>
<td>38/47 (80.9)</td>
</tr>
<tr>
<td>FPR (%)</td>
<td>5/21 (23.8)</td>
<td>9/21 (42.9)</td>
</tr>
<tr>
<td>HR+HER2−</td>
<td>35</td>
<td>26</td>
</tr>
<tr>
<td>NPV (%)</td>
<td>3/13 (23.1)</td>
<td>3/6 (50.0)</td>
</tr>
<tr>
<td>FN (%)</td>
<td>10/31 (32.3)</td>
<td>3/22 (13.6)</td>
</tr>
<tr>
<td>PPV (%)</td>
<td>21/22 (95.5)</td>
<td>19/20 (95.0)</td>
</tr>
<tr>
<td>FPR (%)</td>
<td>1/4 (25.0)</td>
<td>1/4 (25.0)</td>
</tr>
<tr>
<td>HR−HER2+</td>
<td>11</td>
<td>11</td>
</tr>
<tr>
<td>NPV (%)</td>
<td>4/8 (50.0)</td>
<td>4/5 (80.0)</td>
</tr>
<tr>
<td>FN (%)</td>
<td>4/6 (66.7)</td>
<td>1/4 (25.0)</td>
</tr>
<tr>
<td>PPV (%)</td>
<td>2/4 (50.0)</td>
<td>3/6 (50.0)</td>
</tr>
<tr>
<td>FPR (%)</td>
<td>2/6 (33.3)</td>
<td>3/7 (42.9)</td>
</tr>
<tr>
<td>TNBC</td>
<td>21</td>
<td>18</td>
</tr>
<tr>
<td>NPV (%)</td>
<td>6/13 (46.2)</td>
<td>4/4 (100.0)</td>
</tr>
<tr>
<td>FN (%)</td>
<td>7/13 (53.9)</td>
<td>0/11 (0.0)</td>
</tr>
<tr>
<td>PPV (%)</td>
<td>6/8 (75.0)</td>
<td>11/14 (78.6)</td>
</tr>
<tr>
<td>FPR (%)</td>
<td>2/8 (25.0)</td>
<td>3/7 (42.9)</td>
</tr>
</tbody>
</table>

pCR, Pathologic complete response; CBE, clinical breast examination; NPV, negative predictive value; FN, false-negative rate; FPR, false-positive rate; PPV, positive predictive value; HR, hormone receptor; HER2, human epidermal growth factor receptor-2; TNBC, triple-negative breast cancer.
well as a 0% FNR in TNBC (15). Our study also reveals a
4/4 (100%) NPV and 0/11 (0%) FNR by imaging in the
TNBC subset. Although ours, as well as the other, published
results in this setting were retrospective analyses with small
cohorts of patients, the consistent superior results of MRI
cCR predicting pCR along with a low FNR are compelling
for the TNBC subset.

The high NPV and low FNR of MRI should be further
explored for TNBC patients with a prospective NACT trial
that utilizes pre-determined end-points for MRI clinical
response evaluations. If a patient achieves a negative MRI
after completing less than the planned number of NACT
cycles, it may be possible to hold the final cycles. Our trial
revealed a poor NPV for CBE (46.2% in TNBC); thus a CBE
evaluation may be a poor indicator of when to order an MRI
for the patient during NACT. One study reported a NPV of
61.5% in TNBC patients by MG, as well as a NPV of 57.9%
in TNBC patients by US, which is superior to the CBE NPV
found in our study for this subtype (15).

Our study demonstrated that the accuracy of prediction
of pCR vs. non-pCR by CBE varies by breast cancer subtype.
Previously published datasets have shown that MRI most
accurately predicts the amount of residual pathologic disease
in HR-negative subtypes rather than HR-positive subtypes
(11, 16-19). The concept of variation in clinical response
evaluation accuracy by breast cancer subtype is important
for the design of future response-guided NACT trials. A possible
design for trials in this setting would be the integration of a
clinical response assessment after a set number of NACT
cycles in order to group patients into responders vs. non-
responders for subsequent randomized NACT regimens.
Trials including patients with any HR-positive or -negative
subtype and a single clinical response assessment modality,
could incorrectly classify patients as responding or non-
responding due to the varying degree of accuracy of
imaging/CBE. This could occur disproportionately by
subtype, as patients would not only be allocated due to their
inherent sensitivity to the initial NACT, but also due to the
accuracy of imaging/CBE clinical response evaluation of
their subtype.

Our study has certain limitations since it was a single
community oncology practice retrospective data analysis. It
was, however, an inclusive cohort study of all patients treated
with NACT at our site regardless of past medical history,
type/duration of NACT, response to NACT and type of clinical
response evaluation used. The definitions of cCR that were
used by CBE and imaging were generalized to the level of
complete absence of disease versus disease presence in any
form. For the CBE, this is arguably subjective depending on
the oncologist performing the exam with the incalculable factor
of inter-observer discrepancy. An imaging cCR was recorded
for the study if the radiology report specifically stated that there
was no longer presence of malignancy, which may vary by
technical settings of the modality and the individual
radiologist’s review of the images. While our definition of cCR
is less precise than using specific quantitative parameters of
disease measurement by the CBE or imaging modalities, it is
a practical way of exploring the ability to predict pCR vs. non-
pCR in an everyday practice setting.

Conclusion

The accuracy of predicting presence or absence of
pathological residual disease by CBE and imaging varies by
breast cancer subtype. Patients with HR+ breast cancer are
less likely to achieve a pCR with traditional NACT and the
low overall ability to predict pCR in HR+ tumors (≤50% by
either CBE or imaging), based on our results, implies little
incentive for incorporation of imaging as a routine clinical
response evaluation post-NACT in HR+ tumors. Furthermore,
imaging is not better than CBE at predicting the presence of
residual pathologic disease. The use of imaging in HR+ tumors post-NACT may provide little to no
additional value that is not already garnered by performance
of a CBE. This should be taken into account during routine
clinical evaluation for response during NACT, as well as in
the planning and design of future response-guided NACT
clinical trials of HR+ breast cancer patients.

Considering the superior performance of imaging in
TNBC with a 100% NPV and 0% FNR in our study, it may
play a critical role in the prediction of pCR post-NACT. This
finding needs to be further explored through larger
prospective trials in TNBC where it may be feasible to
incorporate an MRI score as a planned secondary end-point
within a randomized controlled trial of NACT regimens. If
the superior accuracy of MRI in predicting pCR is confirmed
in a well-powered prospective trial, it is hypothetically
possible that selected early TNBC complete responders could
avoid unnecessary cycles of additional chemotherapy.

The progression of the therapy of early breast cancer
towards an individualized approach can be applied not only
to the use of personalized therapeutics for specific tumor
traits but, also, towards our interpretation of diagnostics that
would guide the treatment of the disease. Our results
describe the variation in accuracy that exists for imaging and
CBE after NACT and could advance the use of value-based
care with mindful selection of diagnostic modalities
according to breast cancer subtype.

References

1 Rastogi P, Anderson SJ, Bear HD, Geyer CE, Kahlenberg MS,
Robidoux A, Margolese RG, Hoehn JL, Vogel VG, Dakhil SR,
Tamkus D, King KM, Pajon ER, Wright MJ, Robert J, Paik S,
Mamounas EP and Wolmark N: Preoperative chemotherapy:
Updates of National Surgical Adjuvant Breast and Bowel Project
3149-3154, 2011.


