Breast Cancer Subtype as a Predictor of Lymph Node Metastasis according to the SEER Registry

Malcolm D. Mattes, Jay K. Bhatia¹, Daniel Metzger¹, Hani Ashamalla¹, Evangelia Katsoulakis¹

Department of Radiation Oncology, West Virginia University, Morgantown; ¹Department of Radiation Oncology, New York Methodist Hospital, Brooklyn, USA

**Purpose:** Breast cancer subtype correlates with response to systemic therapy and overall survival (OS), but its impact on lymphatic spread is incompletely understood. In this study, we used the Surveillance, Epidemiology, and End Results registry to assess whether the subtype can predict the presence of nodal metastasis or advanced nodal stage in breast cancer. **Methods:** A total of 7,274 eligible patients diagnosed with T1-3 infiltrating ductal carcinoma with known estrogen or progesterone hormone receptor (HR) and human epidermal growth factor receptor 2 (HER2) status, who underwent surgical excision of the primary tumor and pathologic lymph node evaluation, were included in this analysis. Patients were categorized into four breast cancer subtypes: HR+/HER2−; HR+/HER2+; HR−/HER2+; and HR−/HER2−. Binary logistic regression analysis was used to determine whether breast cancer subtype, tumor size, tumor grade, patient race, and patient age at diagnosis are independently predictive of lymph node positivity or advanced nodal stage. The Pearson chi-square test was used to determine whether progesterone receptor (PR) status had an impact on the incidence of lymph node positivity in estrogen receptor (ER) positive patients. **Results:** Independent predictors of nodal positivity included breast cancer subtype (p = 0.040), tumor size (p < 0.001), tumor grade (p < 0.001), and patient age (p < 0.001), whereas only tumor size (p = 0.001), grade (p = 0.001), and patient age (p = 0.005) predicted advanced nodal stage. Triple-negative cancers had a significantly lower risk of nodal positivity than the HR+/HER2− subtype (odds ratio, 0.686; p = 0.004), but no other significant differences between subtypes were observed. There was also no difference in lymph node positivity between PR+ and PR− tumors amongst ER+/HER2− (p = 0.228) or ER+/HER2+ tumors (p = 0.713). **Conclusion:** The HR+/HER2− breast cancer subtype has a higher rate of lymph node involvement at diagnosis than the triple-negative subtype. These findings may play a role in guiding regional management considerations if confirmed in further studies.

**Key Words:** Biological tumor markers, Breast neoplasms, Estrogen receptors

INTRODUCTION

The identification of distinct molecular subtypes of breast cancer has led to an increased understanding of the heterogeneity of the disease [1-3], as these subtypes have been shown to correlate with response to systemic therapy [4,5] and overall survival (OS) [6-8]. The impact of subtype on lymphatic spread and locoregional recurrence, however, is not as well defined. For instance, several studies have reported conflicting results regarding the impact of subtype on the probability of nodal metastasis, and even those studies showing a correlation disagree over which subtypes carry the greatest risk [9-13]. Similarly, while most studies have shown that the human epidermal growth factor receptor 2 (HER2) and basal subtypes have higher locoregional recurrence rates following breast conservation therapy [14-17], it remains unclear which subtype is most likely to benefit from postmastectomy radiation therapy [18-20].

At present, there is a great deal of controversy regarding management of the axillary and regional lymph nodes in breast cancer patients, in relation to both the role of axillary lymph node dissection and regional nodal irradiation after breast-conserving surgery or mastectomy [21-24]. In this epidemiologic study, we used the Surveillance, Epidemiology, and End Results (SEER) registry to assess whether breast can-
cancer subtype independently predicts the presence of nodal metastasis and advanced nodal stage, and also determined which patient- and tumor-related risk factors for lymphatic spread might be used to better guide locoregional management decisions.

**METHODS**

The SEER registry of the National Cancer Institute is a comprehensive source of population-based information for all newly diagnosed cancer patients residing in areas in the United States that participate in the SEER program. In this study we identified all patients from the SEER registry with a newly diagnosed American Joint Committee on Cancer (AJCC) seventh edition [25] stage T1-3 infiltrating ductal carcinoma or mixed infiltrating ductal carcinoma with another breast cancer histology (SEER codes 8521-3/3) with known hormone receptor (HR) and HER2 expression status who underwent lumpectomy or mastectomy including surgical excision of at least one lymph node. Since HER2 status has only been a required data element since 2010, only patients from 2010 to 2011 were included in this analysis.

Clinical information retrieved from the SEER registry included estrogen receptor (ER), progesterone receptor (PR), and HER2 expression status; breast cancer subtype; patient age at diagnosis and race; primary tumor size and grade; number of lymph nodes dissected; number of positive lymph nodes; and AJCC N-stage. ER, PR, and HER2 status were classified in the SEER registry according to immunohistochemistry (IHC) as well as fluorescence in situ hybridization (FISH) testing for HER2. Reporting the results of multigene assays such as Oncotype DX™ in place of IHC was discouraged. Tumors were classified as ER+ or PR+ when 1% or more of the cells stained positive, although no additional information was available on the degree of ER or PR positivity. HER2 IHC was coded with the standard 0–3+ scoring system and considered borderline for 2+ IHC with an equivocal FISH result. Breast cancer subtype is classified in the SEER registry into four mutually exclusive categories: HR+/HER2−; HR+/HER2+; HR−/HER2−; and HR−/HER2− (triple-negative) [26]. HR+ was defined as either ER+ or PR+ or borderline, whereas HR− was defined as both ER− and PR−. Patients with borderline HER2 status were treated as having unknown HER2 status and breast cancer subtype. The rationale for classifying borderline ER and PR status as HR+ has been described previously [26]. It should be noted that because the SEER registry lacks information on Ki-67 and the degree of ER/PR expression, the subtype classification reported by the SEER registry and used in the present study should be considered only an estimation of molecular subtype (e.g., luminal A, luminal B, HER2-overexpressing), albeit an estimation that has frequently been utilized in previous studies [7-10,12,13,15,17-20].

Patients were excluded from analysis if the primary tumor size or extent of invasion was unknown, if the primary tumor invaded the subcutaneous tissue or was attached/fixed to pectoral muscle(s) or underlying tissue, if they were classified as having stage T4 or M1 disease at diagnosis, if they did not have any pathologic lymph node evaluation, if lymph node pathology was unknown, or if they underwent any neoadjuvant therapy prior to surgery. Patients with AJCC nodal stage designation N0(i+) were considered N0 for analysis. Of note, the SEER registry does not distinguish which patients underwent sentinel lymph node biopsy. However, we did exclude all patients who underwent less than an excisional biopsy (e.g., core biopsy or fine-needle aspiration) of lymph node(s) in order to minimize false negative pathologic findings. Other noteworthy pathologic information that is not readily accessible in the SEER registry includes the presence of lymphovascular space invasion, perineural invasion, or extranodal extension.

Descriptive statistics were used to report patient demographics and disease characteristics. The chi-square test was used to determine the impact of PR status on the incidence of lymph node metastasis among ER-positive patients. Multivariate assessment using a binary logistic regression analysis was used to determine which factors were independently associated with lymph node positivity (defined as at least one malignant lymph node) and advanced nodal stage (defined as AJCC stage N2-3, which would categorize these patients as Stage III or locally advanced). Only patients with at least 10 dissected lymph nodes were included in the latter analysis of advanced nodal stage, since this is considered an adequate dissection for breast cancer [25]. Statistical analysis was performed using SPSS version 20.0 (IBM Corp., Armonk, USA). A p-value less than 0.05 was considered statistically significant. This study was approved by the local institutional review board (approval number: 1502604087).

**RESULTS**

A total of 7,274 patients met our criteria for analysis. The mean age at diagnosis was 61.7 ± 12.4 years. A total of 84.1% of patients were Caucasian, whereas 7.9% were African-American, 7.3% were Asian/Pacific Islander, and 0.7% were of other/unknown race. Additional clinical information regarding tumor characteristics is shown in Table 1. The primary tumor was excised by lumpectomy in 3,855 patients (53.0%) and mastectomy in 3,419 patients (47.0%). The mean number
Table 1. Tumor characteristics

| Characteristic | No. of patients (%) |
|---------------|---------------------|
| ER status     |                     |
| Positive      | 6,769 (93.1)        |
| Negative      | 505 (6.9)           |
| PR status     |                     |
| Positive      | 6,025 (82.8)        |
| Negative      | 1,871 (16.3)        |
| Borderline/unknown | 62 (0.8)      |
| HER2 status   |                     |
| Positive      | 612 (8.4)           |
| Negative      | 6,662 (91.6)        |
| Molecular subtype |                   |
| HR+/HER2−     | 6,301 (86.6)        |
| HR+/HER2+     | 502 (6.9)           |
| HR−/HER2+     | 110 (1.5)           |
| HR−/HER2−     | 361 (5.0)           |
| T stage       |                     |
| T1            | 4,888 (67.2)        |
| T2            | 2,081 (28.6)        |
| T3            | 305 (4.2)           |
| Grade         |                     |
| I (well differentiated) | 2,079 (28.6) |
| II (moderately differentiated) | 3,747 (51.5) |
| III (poorly/undifferentiated) | 1,287 (17.7) |
| Unknown       | 161 (2.2)           |
| No. of positive lymph nodes |               |
| 0             | 5,197 (71.4)        |
| 1–3           | 1,531 (21.0)        |
| 4–9           | 353 (4.9)           |
| ≥ 10          | 177 (2.4)           |
| At least 1 (number unknown) | 16 (0.2) |

Table 2. Multivariate analysis for lymph node positivity and advanced nodal stage

| Characteristic | Lymph node positivity | Advanced nodal stage* |
|---------------|-----------------------|-----------------------|
|               | Odds ratio | Standard error | p-value | Odds ratio | Standard error | p-value |
| Race          |            |                |         |            |                |         |
| White         | Reference   | 0.081           | <0.001  | Reference   | 1.013           | 0.004   | 0.005 |
| Asian         | 0.99        | 0.109           | 0.287   | 0.95        | 0.208           | 0.807   |
| Black         | 1.254       | 0.1             | 0.024   | 0.967       | 0.19             | 0.859   |
| Tumor size    | 1.052       | 0.002           | <0.001  | 1.032       | 0.003           | <0.001  |
| Tumor grade   | 1.433       | 0.045           | <0.001  | 1.357       | 0.09             | 0.001   |
| Cancer subtype|            |                |         |            |                |         |
| HR+/HER2−     | Reference   | 0.040           | 0.439   |
| HR+/HER2+     | 0.927       | 0.109           | 0.491   | 1.009       | 0.198           | 0.966   |
| HR−/HER2+     | 0.928       | 0.233           | 0.739   | 1.355       | 0.331           | 0.359   |
| HR−/HER2−     | 0.686       | 0.132           | 0.004   | 0.744       | 0.244           | 0.226   |

HR= hormone receptor; HER2= human epidermal growth factor receptor 2; HR= hormone receptor.

only a trend toward significance for patient race as a whole (p = 0.081). Older patient age (p = 0.005), larger tumor size (p < 0.001), and higher tumor grade (p = 0.001) were independent predictors of advanced nodal stage. Finally, there was no significant difference in the incidence of lymph node positivity between PR+ and PR− tumors within the groups of ER+/HER2− (25.6% vs. 28.1%, respectively, p = 0.228) or ER+/HER2+ tumors (34.6% vs. 32.7%, respectively, p = 0.713).

DISCUSSION

In this large cohort of patients, we have shown that in addition to traditional risk factors for lymphatic spread, such as tumor size, tumor grade, and patient age, the HR+/HER2− subtype of breast cancer also independently predicts a higher rate of nodal involvement with malignancy compared to the triple-negative subtype. In fact, tumor grade and breast cancer subtype appear to have the greatest magnitude of effect according to the reported odds ratios (ORs). While we did not observe a similar impact of subtype in predicting advanced nodal stage, the OR was similar between HR+/HER2− and triple-negative patients, and it is possible that the smaller number of patients in this analysis may have limited its power to detect a difference.

Several single-institution series have previously addressed the impact of breast cancer subtype on lymphatic spread, although their findings are heterogeneous. Jones et al. [9] found that breast cancer subtype had no association with nodal positivity, N stage, or the absolute number of nodes involved among 453 patients who underwent breast-conserving sur-
gery for stage I-II breast cancer. On the other hand, Reyal et al. [10] showed in a series of 2,654 patients that the triple-negative subtype had the lowest incidence of sentinel lymph node positivity and the HER2 subtype had the highest incidence of sentinel lymph node positivity. Similarly, Crabb et al. [11] demonstrated in a series of 4,444 patients that the triple-negative subtype had the lowest risk of lymph node metastasis. Zhou et al. [12] also reported that among 130 patients with positive sentinel lymph nodes, the luminal A and B subtypes had a higher likelihood of having additional nodal metastases upon axillary lymph node dissection than the triple-negative or HER2 subtypes. Interpreting such disparate findings in the context of our own results is challenging. The difference we observed between HR+/HER2− and triple-negative tumors reiterates the findings of the majority of the aforementioned publications. However, our reported lack of difference between HR−/HER2+ tumors and other subtypes may have been observed because of a lack of power to detect a difference in the small subgroup of patients with this subtype in the SEER registry. Other discrepancies may be due to the fact that in all of these studies breast cancer subtype was only approximated by immunohistochemical analysis, which is commonly used as a surrogate for molecular gene expression profiling, but may not be as accurate in determining molecular subtype due to differences in institutional protocols and techniques.

Another approach to assess the value of various regional management approaches is to assess patterns of recurrence. Unfortunately the SEER database does not provide information regarding recurrence, and therefore recurrence rates were not evaluated in this study. While the available literature generally supports the triple-negative and HER2 subtypes as having higher locoregional recurrence rates after breast conservation therapy and mastectomy [14–17], several studies suggest that postmastectomy radiation therapy is actually most beneficial to the luminal A and B subtypes [18–20], perhaps reflecting the fact that the basal and HER2 subtypes have a higher rate of distant spread even in the absence of lymphatic spread [27], thus making locoregional therapy relatively less important to the OS of these patients. On the other hand, there is also evidence from a randomized trial by Wang et al. [27] that node-negative, triple-negative patients do have a survival benefit associated with postmastectomy radiation, perhaps reflecting the importance of the patient population, the presence of other prognostic factors, and variability in treatment techniques in interpreting these outcomes data.

PR expression in ER-positive patients has been shown to be predictive of response to Tamoxifen and Oncotype DX™ recurrence score and prognostic of breast cancer-specific survival [28,29]. One rationale used to explain these findings is that the ER-positive, PR-negative phenotype potentially represents a de-differentiated and more aggressive tumor. Van Calster et al. [13] showed in a series of 2,227 patients that among ER+/HER2+ tumors (but not ER−/HER2+ tumors), PR positivity was associated with a higher rate of lymph node involvement than that found in PR-negative tumors. Our findings in a larger cohort of patients suggest that PR expression does not impact lymphatic spread in any ER-positive patients.

The main strength of using the SEER registry is that it provides access to a larger cohort of patients than is generally feasible at single institutions. However, there remain several limitations as well. First, any SEER study of this nature is limited by the lack of a central pathology review and uniformity of laboratory techniques. The power of our findings was also limited by the types of patients present in SEER participating areas, who were predominantly white and had ER+ HER2− tumors. The lower percentage of HER2+ tumors in this study compared to the overall SEER incidence may also have been a consequence of excluding patients who had metastasis at diagnosis or had received neoadjuvant chemotherapy [26], though excluding these patients is unlikely to have affected our analysis of T1-3 breast cancer and was necessary to ensure a relatively homogeneous cohort of patients without any previous treatments that could confound pathologic interpretation of nodal involvement with malignancy. We also had no information on whether any of the lymph node(s) examined were sentinel lymph nodes, and although sentinel lymph node biopsy has long been the standard of care in the United States and is probably the most common technique used in this contemporary cohort of patients in cases in which only a small number of dissected lymph nodes was reported, we have no way to quantify what percentage did not undergo sentinel lymph node biopsy. It is also noteworthy that we were not able to obtain information on certain factors such as lymphovascular invasion, which has been associated with lymph node metastasis in previous studies [23,24]. As such, there may exist additional factors that correlate with lymphatic spread which we could not evaluate in this study.

In conclusion, according to the SEER registry, breast cancer subtype is an independent risk factor for lymph node positivity, with the HR+/HER2− subtype carrying a greater risk of nodal involvement with malignancy than the triple-negative subtype. This suggests that additional methods of regional lymph node diagnosis and management, such as axillary lymph node dissection beyond a sentinel lymph node and lymph node directed radiation therapy, may have less value in patients with triple-negative tumors, even though they are known to carry a worse prognosis. Additional predictive factors identified in this as well as other studies, such as tumor
size, grade, and patient age, should help guide management in these patients.

**CONFLICT OF INTEREST**

The authors declare that they have no competing interests.

**REFERENCES**

1. Perou CM, Sorlie T, Eisen MB, van de Rijn M, Jeffrey SS, Rees CA, et al. Molecular portraits of human breast tumours. Nature 2000;406:747-52.

2. Sotiriou C, Neo SY, McShane LM, Korn EL, Long PM, Jazaeri A, et al. Breast cancer classification and prognosis based on gene expression profiles from a population-based study. Proc Natl Acad Sci U S A 2003;100:10933-8.

3. Sorlie T, Perou CM, Tibshirani R, Aas T, Geisler S, Johnsen H, et al. Gene expression patterns of breast carcinomas distinguish tumor subclasses with clinical implications. Proc Natl Acad Sci U S A 2001;98:10861-9.

4. Brenton JD, Carey LA, Ahmed AA, Caldas C. Molecular classification and molecular forecasting of breast cancer: ready for clinical application? J Clin Oncol 2005;23:7350-60.

5. Rouzier R, Perou CM, Symmans WF, Ibrahim N, Cristofanilli M, An DD, et al. Basal breast cancer molecular subtype predicts for lower incidence of axillary lymph node metastases: the American College of Surgeons Oncology Group Z0011 randomized trial. Ann Surg Oncol 2013;20:2866-72.

6. Rouzier R, Perou CM, Symmans WF, Ibrahim N, Cristofanilli M, An DD, et al. Basal breast cancer molecular subtypes respond differently to preoperative chemotherapy. Clin Cancer Res 2005;11:5678-85.

7. Reis-Filho JS, Pusztai L. Gene expression profiling in breast cancer: classification, prognostication, and prediction. Lancet 2011;378:1812-23.

8. Hugh J, Hanson J, Cheang MC, Nielsen TO, Perou CM, Dumontet C, et al. Breast cancer subtypes and response to docetaxel in node-positive breast cancer: use of an immunohistochemical definition in the BCIRG 001 trial. J Clin Oncol 2009;27:1168-76.

9. Carey LA, Perou CM, Livasy CA, Dressler LG, Cowan D, Conway K, et al. Race, breast cancer subtypes, and survival in the Carolina Breast Cancer Study. JAMA 2006;295:2492-502.

10. Jones T, Neboor H, Wu H, Yang Q, Haffty BG, Evans S, et al. Are breast cancer subtypes prognostic for nodal involvement and associated with clinicopathologic features at presentation in early-stage breast cancer? Ann Surg Oncol 2013;20:2866-72.

11. Seyal F, Rouzier R, Depont-Hazdelt B, Bollet MA, Pierga JY, Alran S, et al. The molecular subtype classification is a determinant of sentinel node positivity in early breast carcinoma. PLoS One 2011;6:e20297.

12. Voduc KD, Cheang MC, Tykdesley S, Gelmon K, Nielsen TO, Kennecke H. Breast cancer subtypes and the risk of local and regional relapse. J Clin Oncol 2010;28:1684-91.

13. Wo JY, Taghian AG, Nguyen PL, Raad RA, Sreedhara M, Bellon JR, et al. The association between biological subtype and isolated regional nodal failure after breast-conserving therapy. Int J Radiat Oncol Biol Phys 2010;77:188-96.

14. Voduc KD, Cheang MC, Tykdesley S, Gelmon K, Nielsen TO, Kennecke H. Breast cancer subtypes and the risk of local and regional relapse. J Clin Oncol 2010;28:1684-91.

15. Wo JY, Taghian AG, Nguyen PL, Raad RA, Sreedhara M, Bellon JR, et al. The association between biological subtype and isolated regional nodal failure after breast-conserving therapy. Int J Radiat Oncol Biol Phys 2010;77:188-96.

16. Millar EK, Graham PH, O’Toole SA, McNeil CM, Browne L, Morey AL, et al. Prediction of local recurrence, distant metastases, and death after breast-conserving therapy in early-stage invasive breast cancer using a five-biomarker panel. J Clin Oncol 2009;27:4701-8.

17. Nguyen PL, Taghian AG, Katz MS, Niemierko A, Abi Raad RF, Boon WL, et al. Breast cancer subtype approximated by estrogen receptor, progesterone receptor, and HER-2 is associated with local and distant recurrence after breast-conserving therapy. J Clin Oncol 2008;26:2373-8.

18. Wu SG, He ZY, Li Q, Li FY, Lin Q, Lin HX, et al. Predictive value of breast cancer molecular subtypes in Chinese patients with four or more positive nodes after postmastectomy radiotherapy. Breast 2012;21:657-61.

19. Kyndi M, Sorensen FB, Knudsen H, Overgaard M, Nielsen HM, Overgaard J, et al. Estrogen receptor, progesterone receptor, HER-2, and response to postmastectomy radiotherapy in high-risk breast cancer: the Danish Breast Cancer Cooperative Group. J Clin Oncol 2008;26:1419-26.

20. Dominici LS, Mittendorf EA, Wang X, Liu J, Kuerer HM, Hunt KK, et al. Implications of constructed biologic subtype and its relationship to locoregional recurrence following mastectomy. Breast Cancer Res 2012;14:R82.

21. Giuliano AE, McCall L, Beitsch P, Whitworth PW, Blumencranz P, Leitch AM, et al. Locoregional recurrence after sentinel lymph node dissection with or without axillary dissection in patients with sentinel lymph node metastases: the American College of Surgeons Oncology Group Z0011 randomized trial. Ann Surg 2010;252:426-32.

22. Whelan TJ, Oliverio I, Ackerman I, Chapman JW, Chua B, Nabil A, et al. NCIC-CTG MA.20: an intergroup trial of regional nodal irradiation in early breast cancer. 2011 ASCO Annual Meeting. 2011;29. Abstract #LBA1003.

23. Recht A, Gray R, Davidson NE, Fowble BL, Solin LJ, Cummings FJ, et al. Locoregional failure 10 years after mastectomy and adjuvant chemotherapy with or without tamoxifen without irradiation: experience of the Eastern Cooperative Oncology Group. J Clin Oncol 1999;17:1689-700.

24. Taghian A, Jeong JH, Mamounas E, Anderson S, Bryant J, Deutsch M, et al. Patterns of locoregional failure in patients with operable breast cancer treated by mastectomy and adjuvant chemotherapy with or without tamoxifen and without radiotherapy: results from five National Surgical Adjuvant Breast and Bowel Project randomized clinical trials. J Clin Oncol 2004;22:4247-54.

25. Edge SB, Compton CC. The American Joint Committee on Cancer: the 7th edition of the AJCC cancer staging manual and the future of TNM. Ann Surg Oncol 2010;17:1471-4.

26. Howlader N, Altekruse SF, Li CI, Chen VW, Clarke CA, Ries LA, et al. US incidence of breast cancer subtypes defined by joint hormone receptor and HER2 status. J Natl Cancer Inst 2014;106:1055.

27. Wang J, Shi M, Ling R, Xia Y, Luo S, Fu X, et al. Adjuvant chemotherapy
and radiotherapy in triple-negative breast carcinoma: a prospective randomized controlled multi-center trial. Radiother Oncol 2011;100:200-4.

28. Mattes MD, Mann JM, Ashamalla H, Tejwani A. Routine histopathologic characteristics can predict oncotype DX(TM) recurrence score in subsets of breast cancer patients. Cancer Invest 2013;31:604-6.

29. Liu S, Chia SK, Mehl E, Leung S, Raiput A, Cheang MC, et al. Progesterone receptor is a significant factor associated with clinical outcomes and effect of adjuvant tamoxifen therapy in breast cancer patients. Breast Cancer Res Treat 2010;119:53-61.