

RESEARCH ARTICLE

A Breast Cancer Nomogram for Prediction of Non-Sentinel Node Metastasis - Validation of Fourteen Existing Models

Bulent Koca¹, Bekir Kuru^{2*}, Necati Ozen², Savas Yoruker², Yuksel Bek³

Abstract

Background: To avoid performing axillary lymph node dissection (ALND) for non-sentinel lymph node (SLN)-negative patients with-SLN positive axilla, nomograms for predicting the status have been developed in many centers. We created a new nomogram predicting non-SLN metastasis in SLN-positive patients with invasive breast cancer and evaluated 14 existing breast cancer models in our patient group. **Materials and Methods:** Two hundred and thirty seven invasive breast cancer patients with SLN metastases who underwent ALND were included in the study. Based on independent predictive factors for non-SLN metastasis identified by logistic regression analysis, we developed a new nomogram. Receiver operating characteristics (ROC) curves for the models were created and the areas under the curves (AUC) were computed. **Results:** In a multivariate analysis, tumor size, presence of lymphovascular invasion, extranodal extension of SLN, large size of metastatic SLN, the number of negative SLNs, and multifocality were found to be independent predictive factors for non-SLN metastasis. The AUC was found to be 0.87, and calibration was good for the present Ondokuz Mayıs nomogram. Among the 14 validated models, the MSKCC, Stanford, Turkish, MD Anderson, MOU (Masaryk), Ljubljana, and DEU models yielded excellent AUC values of > 0.80. **Conclusions:** We present a new model to predict the likelihood of non-SLN metastasis. Each clinic should determine and use the most suitable nomogram or should create their own nomograms for the prediction of non- SLN metastasis.

Keywords: Breast cancer nomogram - non-sentinel node metastasis - nomogram for non-SLN metastasis

Asian Pac J Cancer Prev, 15 (3), 1481-1488

Introduction

Axillary lymph node dissection (ALND), which has long been used to evaluate the status of the axillary nodes, is known to be associated with a high morbidity rate (Kell et al., 2010). Sentinel lymph node biopsy (SLNB) has successfully replaced ALND for evaluation of the axilla in breast cancer patients with clinically negative axilla (Kell et al., 2010). However, ALND performed in SLNB-positive patients demonstrated the absence of metastasis in 22-71% of the non-SLNs (Zhu et al., 2013).

To avoid performing ALND for non-SLN-negative patients with SLN-positive axilla, nomograms and scoring systems for predicting the status of the axillary non-SLNs have been developed in many centers (Van Zee et al., 2003; Degnim et al., 2005; Smidt et al., 2005; Chapgar et al., 2006; Pal et al., 2008; Alran et al., 2007; Kohrt et al., 2008; Cho et al., 2008; Coufal et al., 2009; Coutant et al., 2009; Gur et al., 2010; Perhavec et al., 2010; Lombardi, et al., 2011; Mittendorf et al., 2012; Meretoja et al., 2012; Derici et al., 2012). Many studies have tested these models and have obtained varying results (Van Zee et al., 2003; Kocsis et al., 2004; Degnim et al., 2005; Soni et al.,

2005; Lambert et al., 2006; Ponzzone et al., 2007; Pal et al., 2008; Kohrt et al., 2008; Cho et al., 2008; Klar et al., 2008; Poirier et al., 2008; Coufal et al., 2009; Coutant et al., 2009; Coutant et al., 2009; Scow et al., 2009; Gur et al., 2010; Moghaddam et al., 2010; Hessman et al., 2011; Hidar et al., 2011; Lombardi et al., 2011; Tan et al., 2011; Derici et al., 2012; Chen et al., 2012; Zhu et al., 2012; Sasada et al., 2013).

The characteristics of the breast cancer patients varied across different centers (Pal et al., 2008; Coutant et al., 2009). Therefore, before employing a given model, each center should validate the nomograms or scoring systems and use the best suitable model, or if possible, each center should create their own nomograms. However, the ACOSOG Z0011 study proved that among patients who underwent breast-conserving therapy, locoregional recurrence and survival with and without completion ALND were not significantly different for clinical T1-2 tumors with 1-2 positive SLNs without extranodal extension (Giuliano et al., 2010). For this subgroup of patients, the use of nomograms could be unnecessary, but for patients beyond the scope of the ACOSOG Z0011 study and for patients who are candidates for mastectomy,

¹General Surgery, Korgan Devlet Hastanesi, Ordu, ²Department of General Surgery, ³Department of Biostatistics and Medical Information, Ondokuz Mayıs University School of Medicine, Samsun, Turkey *For correspondence: bekirkuru@gmail.com, bekirkuru@yahoo.com

it may still be important to predict the status of the non-SLNs. To the best of our knowledge, the present study is also the first to validate the Rome and European models other than by using the datasets reported in the original studies (Lombardi et al., 2011; Derici et al., 2012).

Our aims were to investigate the factors that predict non-SLN metastasis, to create a nomogram, and to validate the 14 existing models in our patient group.

Materials and Methods

Two hundred and thirty seven invasive breast cancer patients with positive SLNB who underwent ALND between 2003 and 2012 in the Department of General Surgery at Ondokuz Mayıs University School of Medicine were included in the study. The subjects included in the present study were selected from 739 invasive breast cancer patients with T1-3 tumors and clinically negative axilla who had not received neoadjuvant chemotherapy and who underwent SLNB and breast-conserving surgery or mastectomy. The study design was approved by the Ondokuz Mayıs University Medical Research and Ethical Board.

The probable predictive factors for non-SLN metastasis were defined as: age (<50, ≥50), pathological tumor size (pT) (≤1 cm, 1.1-2 cm, 2.1-3 cm, 3.1-5 cm, >5 cm), histological type (invasive ductal, invasive lobular), histological grade (ductal 1, ductal 2, ductal 3, lobular), lymphovascular invasion (LVI) (absent, present), number of positive SLNs (posSLN), number of negative SLNs (negSLN), ratio of positive SLNs to total SLNs (pRatio) (<0.5, ≥0.5-1, 1), the size of the largest SLN metastasis in mm (SLNMS), multifocality (absent, present), extranodal extension of the SLN (ENE) (absent, present), the Estrogen receptor (ER) status (negative, positive), the Progesterone receptor (PR) status (negative, positive) and the c-erb B2 receptor status (negative, positive) (Table 1).

Patients underwent sentinel lymph node biopsy with 5 mL injectable sterile solutions of 1% isosulphan blue. Patients with SLN metastases in frozen sections underwent immediate ALND, and patients found to have SLN metastasis by routine or serial section H&E later underwent a second surgery for ALND. All sentinel lymph nodes were sent for frozen analysis. If the SLN size was ≤1 cm, it was bisected parallel to the long axis. An imprint was applied for two cut surfaces, which then underwent frozen sectioning. The frozen sections and imprint prepartes were stained with H&E and analyzed under a microscope. If the SLN >1 cm, it was cut into slices perpendicular to the long axis at 3 mm intervals. All cut surfaces underwent imprint and frozen analyses. If the SLN contained apparent metastases at the macroscopic evaluation, only the imprint analysis was performed. After the frozen section analysis, the remaining frozen tissue was fixed in formalin and embedded in paraffin for routine pathological examination. The non-SLNs that were ≤1 cm obtained from ALND were bisected parallel to the long axis, and the non-SLNs that were >1 cm were cut into slices perpendicular to the long axis at 3 mm intervals. The evaluation of non-SLNs was usually performed only by H&E, and serial sectioning or immunohistochemistry

were not routinely performed.

The factors that were found to be associated with non-SLN metastasis were entered into the logistic regression analysis. Independent predictive factors and predictive probabilities of non-SLN metastasis were determined by backward logistic regression analysis. Receiver operating characteristic (ROC) curves were generated for the calculations of the area under the curve (AUC). We used a logistic regression modeling approach to describe the relationship between several predictive factors (pT, LVI, ENE, negSLN, SLNMS, and multifocality) and the expected (E) value of the dichotomous dependent variable Y (non-SLN metastasis). The formula for the logistic model to describe the probability of occurrence for the outcome of Y is as follows (Kleinbaum et al., 1998): $E(Y)=pr(Y=1)=1/[1+e^{-(\beta_0+\sum_j\beta_jX_j)}]$. Then, by transformation of the probability $pr(Y=1)$, the logit form of the model emerges as $\text{logit}[pr(Y=1)]=\beta_0+\sum_j\beta_jX_j$. Using the logistic model $P=1/[1+e^{-(\beta_0+\sum_j\beta_jX_j)}]$, we created the prediction equation for the probability of non-SLN metastasis (P) as $P=1/[1+e^{-(2.746+0.45*pT+1.242*LVI+0.896*ENE+(-0.67*negSLN)+0.095*SLNMS+1.227*multifocality)}]$.

We validated 14 models that had over 100 patients with SLN metastasis. For calculation of the likelihood of non-SLN metastasis according to the MSKCC, Stanford and MD Anderson nomograms for the present series, we used the formulas available for use on the websites of MSKCC (<http://nomograms.mskcc.org/Breast/BreastAdditionalNonSLNMetastasesPage.aspx>), Stanford University (<https://www3-hrpdc.stanford.edu/nsln-calculator/>), and MD Anderson Center (http://www3.mdanderson.org/app/medcalc/bc_nomogram2/index.cfm?pagename=nsln). For the other models, we used the reported formulas or criteria described in the relevant articles. For the 14 previously reported models and our new Ondokuz Mayıs nomogram, Receiver Operating Characteristic (ROC) curves were generated, and the areas under the curve (AUC) were calculated to assess the discrimination of the models. The AUC varies between 0.5 and 1.0, and a higher value is better. Discrimination which refers to the ability of a model to distinguish patients at high risk for positive non-SLN from low risk patients was quantified by AUC (Hanley and McNeil, 1982).

Scatter plots were generated to assess the agreement of the model-predicted probabilities and were evaluated visually. The calibration (i.e., the ability of a predictive model to match the predicted and observed probabilities) was assessed graphically. The predicted probabilities were categorized into 10 deciles, and the observed percentages of positive non-SLNs for each decile (actual probability) were calculated. Using the actual probability as the Y-axis and the mean predicted probability as the X-axis, the calibration curve was generated. The number and proportion of patients and the false negative rates at various scores or cut-off levels were also calculated for the Ondokuz Mayıs nomogram and for each model validated using our patient series. The false negative rates were estimated as the number of patients with non-SLN metastasis by the number of patients at the scores or cut-off values.

The factors for comparison were recorded on a

computer using Statistical Program for Social Science (SPSS) version 15.0. The categorical data were expressed as numbers and percentages, and the continuous data were expressed as the mean±standard deviation or median (range). Comparisons of positive non-SLNs with categorical data were performed using chi-square tests, and comparisons with numerical data were performed using Mann-Whitney-U tests. A $p < 0.05$ was accepted as the significance level. Multivariate analysis was performed using logistic regression analysis and odds ratios (OR) for positive non-SLNs, and the 95% confidence intervals (CI) were calculated. The calibration of the Ondokuz Mayıs nomogram was evaluated using the Hosmer-Lemeshow goodness-of-fit test and visually by plots (Hosmer and Lemeshow, 2006).

Results

The median age was 49 (range, 26-78), and the median number of axillary nodes removed was 14 (range, 7-45). The median number of SLNs and metastatic SLNs were 2 (range, 1-8) and 1 (range, 1-7), respectively, and the median number of removed non-SLNs and metastatic non-SLNs were 11 (0-41) and 3 (1-22), respectively. The median size of the SLN metastases was 9 mm (range, 1-25). While 115 patients (48.5%) out of 237 patients with SLN metastasis had non-SLN metastasis, 122 (51.5%) did not. One hundred and twenty seven patients (54%) underwent breast-conserving surgery (BCS) and 110 (46%) mastectomy. Two hundred and twenty six (95%) SLN metastases were detected in the frozen sections, 7 at routine H&E, and 5 by immunohistochemistry (IHC).

By univariate analysis, pT ($p < 0.001$), LVI ($p < 0.001$), SLNMS ($p < 0.001$), ENE ($p < 0.001$), negSLN ($p < 0.001$), pRatio ($p < 0.001$), c-erb B2 receptor ($p = 0.037$), and multifocality ($p < 0.001$) were found to be significantly associated with non-SLN metastasis (Table 1). PosSLN was also associated with non-SLN metastasis ($p = 0.054$). Multivariable logistic regression analysis demonstrated that pT (OR 1.6, 95%CI 1.10-2.29), presence of LVI (OR 3.5, 95%CI 1.69-7.06), presence of ENE (OR 2.4, 95%CI 1.09-5.48), negSLN (OR 0.5, 95%CI 0.36-0.72), SLNMS (OR 1.1, 95%CI 1.02-1.19), and multifocality (OR 3.4, 95%CI 1.06-10.92) were found to be independent factors predicting non-SLN metastasis (Table 2).

The generated ROC curves for the present series and for the 14 other models are presented in Figures 1 and 2. The calculated AUC value for the Ondokuz Mayıs nomogram was 0.871 (CI 0.82-0.91). The AUC values, sensitivity, specificity, false positive (FP) and false negative rates (FN) and the negative and positive predictive values (NPV and PPV) for the Ondokuz Mayıs nomogram and for the other 14 models validated using our dataset (and the number of patients and the AUC values reported in the original studies) are presented in Table 3. The calibration plot for the Ondokuz Mayıs model is presented in Figure 3. The Hosmer-Lemeshow goodness-of-fit test revealed that the P-value was 0.18, suggesting a good calibration. The agreement of the predicted probabilities of the Ondokuz Mayıs model with those of the MD Anderson and European nomograms

Table 1. Comparison of Patient and Tumor Characteristics by Positive Non-SLNs

Characteristics	No. (%)	No. of positive non-SLN (%)	p univariate	p multivariate
Age	≤50	132 (56)	0.89	--
	>50	105 (44)		
Pathological tumor size (pT)	≤1 cm	26 (11)	<0.001	0.009
	1.1-2 cm	62 (26)		
	2.1-3 cm	68 (29)		
	3.1-5 cm	66 (28)		
	>5 cm	15 (6)		
Histological type	Invasive ductal	213 (90)	1.0	--
	Invasive lobular	24 (10)		
Tumor type and histological grade	Ductal 1	18 (8)	0.47	--
	Ductal 2	136 (57)		
	Ductal 3	59 (25)		
	Lobular	24 (10)		
	Lobular	12 (50)		
Lymphovascular invasion (LVI)	Absent	97 (41)	<0.001	0.001
	Present	140 (59)		
Extranodal extension (ENE)	Absent	169 (71)	<0.001	0.037
	Present	68 (29)		
Multifocality	Absent	205 (86)	<0.001	0.043
	Present	32 (14)		
Estrogen receptor	Negative	29 (12)	0.32	--
	Positive	208 (88)		
Progesteron receptor	Negative	73 (31)	0.48	--
	Positive	164 (69)		
c-erb B2 receptor	Negative	176 (74)	0.037	NS
	Positive	61 (26)		
pRatio	<0.5	40 (17)	<0.001	NS
	≥0.5-<1	86 (36)		
	1	111 (47)		
Number of positive SLNs (posSLN)	1	130 (55)	0.054	NS
	2	79 (33)		
	3	9 (4)		
	4	9 (4)		
	5	6 (2)		
	6	2 (1)		
	7	2 (1)		
Number of negative SLNs (negSLN)	0	109 (46)	<0.001	<0.001
	1	76 (32)		
	2	30 (13)		
	3	19 (8)		
	4	3 (1)		
SLN largest metastasis size in mm (SLNMS) Median (range)	9 (1-25)	----	<0.001	0.019
SLN metastasis size	≤2 mm	30 (13)	<0.001	NA
	> 2 mm	207 (87)		

*pRatio: Ratio of metastatic SLNs to total SLNs, SLN: sentinel lymph node, NA: not assessed

Table 2. Multivariable logistic-regression analysis for non-SLN metastasis

Factors	OR	95% CI	P
Lymphovascular invasion (LVI)	3.5	1.69-7.06	0.001
SLN largest metastasis size in mm (SLNMS)	1.1	1.02-1.19	0.019
Tumor size (pT)	1.6	1.10-2.29	0.012
Number of negative SLNs (negSLN)	0.5	0.36-0.72	<0.001
Extranodal extension (ENE)	2.4	1.09-5.48	0.029
Multifocality	3.4	1.06-10.92	0.039

SLN: sentinel lymph node

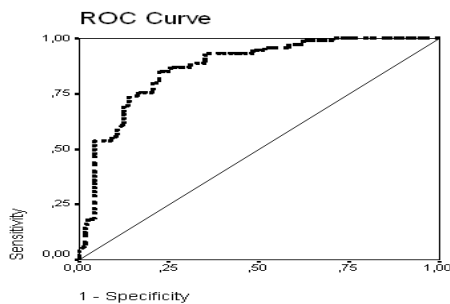


Figure 1. ROC Curve and AUC for the Ondokuz Mayıs (OMU) Nomogram

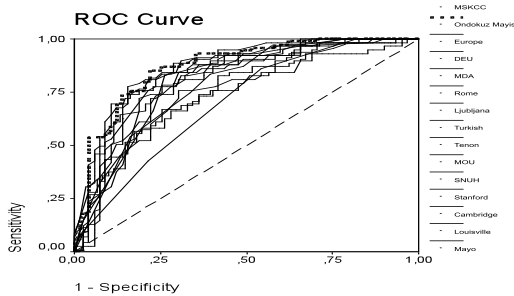


Figure 2. ROC Curves and AUCs for the Ondokuz Mayıs and 14 Other Models

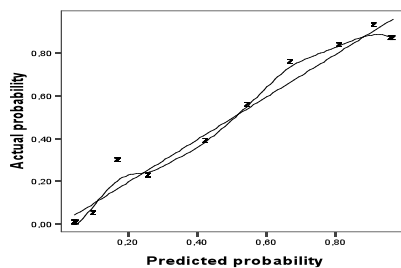


Figure 3. Calibration Plot of the Ondokuz Mayıs Nomogram

were assessed by Scatter plots (data not shown). The MD Anderson model was found to have high agreement with the Ondokuz Mayıs model, whereas the MSKCC, Turkish, MOU (Masaryk), Ljubljana, and DEU models showed modest agreement. The Louisville, Tenon, SNUH (Seoul), European, Mayo, Rome, Cambridge, and Stanford models were found to exhibit low agreement, with wide scatter around the line of agreement. Our findings demonstrated that the MSKCC, Stanford, Turkish, MOU (Masaryk), Ljubljana, MD Anderson, and DEU models can effectively predict non-SLN metastasis.

To assess the clinical utility of the nomograms, false negative rates estimated at various scores or cut-off levels were evaluated to define a subgroup of patients with a low predicted probability of non-SLN metastasis. Coutant et al. (Coutant et al., 2009) and Mittendorf et al. (Mittendorf et al., 2012) estimated the FN rates among patients with a predictive probability of ≤ 10 for clinical utility. This is a useful choice for clinical use because the false negative rate of SLN is usually accepted as $< 10\%$ (Kohrt et al., 2008; Poirier et al., 2008; Coufal et al., 2009; Hidar et al., 2011). However, the predictive probability levels of $\leq 15\%$ and $\leq 20\%$ have also been reported to be accepted as a definition of a subgroup with a low predicted probability of non-SLN metastasis to avoid ALND or to

Table 4. Number of Patients and False Negative Rates at Various Scores or Cut-off values by Models for 237 Patients with Positive SLNs, and Ondokuz Mayıs Nomogram

Model N=237	AUC	Cut-off value or score	No. of* Patients (%)	No. of patients with non-SLN metastasis	False negative **Rates %
MSKCC	0.84	$\leq 10\%$	10 (4)	0	0
		$\leq 15\%$	12 (5)	0	0
		$\leq 20\%$	14 (6)	2	14
Mayo Clinic	0.73	$\leq 10\%$	16 (7)	4	25
		$\leq 15\%$	17 (7)	5	29
		$\leq 20\%$	18 (8)	6	33
Louisville	0.70	1	4 (2)	0	0
Cambridge	0.80	$\leq 10\%$	8 (3,4)	0	0
		$\leq 15\%$	21 (9)	0	0
		$\leq 20\%$	36 (15)	0	0
Stanford	0.81	$\leq 10\%$	15 (6)	0	0
		$\leq 15\%$	25 (11)	1	4
		$\leq 20\%$	37 (16)	2	5.4
SNUH (Seoul)	0.80	≤ 0.5	24 (10)	0	0
		≤ 1	46 (19)	4	9
		$\leq 15\%$	31 (13)	0	0
MOU (Masaryk)	0.83	$\leq 10\%$	16 (7)	0	0
		$\leq 15\%$	47 (20)	5	11
		$\leq 20\%$	38 (16)	2	5.3
Tenon	0.78	≤ 3	50 (21)	4	8
		≤ 3.5	14 (6)	0	0
		$\leq 15\%$	16 (7)	0	0
Turkish	0.82	$\leq 10\%$	36 (15)	4	11
		$\leq 15\%$	8 (3)	0	0
		$\leq 20\%$	12 (5)	2	17
Ljubljana	0.81	$\leq 10\%$	0	--	--
		$\leq 15\%$	4 (2)	0	0
		$\leq 20\%$	13 (5)	0	0
Rome	0.75	$\leq 10\%$	8 (3)	0	0
		$\leq 15\%$	8 (3)	0	0
		$\leq 20\%$	12 (5)	2	17
MD Anderson	0.85	$\leq 10\%$	10 (4)	0	0
		$\leq 15\%$	28 (12)	0	0
		$\leq 20\%$	36 (15)	0	0
DEU (9 Eylül)	0.84	$\leq 10\%$	75 (32)	8	11
		$\leq 15\%$	82 (35)	11	13
		$\leq 20\%$	109 (46)	19	17
European	0.75	$\leq 10\%$	90 (38)	25	28
		$\leq 15\%$	122 (51)	38	31
		$\leq 20\%$	153 (65)	53	35
Ondokuz Mayıs	0.87	$\leq 10\%$	36 (15)	0	0
		$\leq 15\%$	55 (23)	4	7
		$\leq 20\%$	67 (28)	6	9

*No. of patients with equal or lower than scores or predicted probabilities (cut-off values). FNR: false negative rates; **Estimated by the patients with non-SLN metastasis among No. of the patients

be without axillary recurrence in the absence of ALND (Lambert et al., 2006; Ponzzone et al., 2007; Zakarai et al., 2008; Hidar et al., 2011). Therefore, we also investigated the FN rates and the proportion of patients with $\leq 10\%$, $\leq 15\%$, and $\leq 20\%$ cut-off levels of predicted probabilities obtained from the models (Table 4). While in the Mayo and European models, the FN rates at the $\leq 10\%$ cut-off level were 25% and 28% and were 33% and 35% at the $\leq 20\%$ cut-off level, respectively; in the Ondokuz Mayıs nomogram, the FN rates at the $\leq 10\%$ and $\leq 20\%$ cut-off levels were 0% and 9%, respectively. The median predictive probabilities in the SLN metastasis only group and the non-SLN metastasis group based on the present Ondokuz Mayıs model were 20% and 80%, respectively..

Table 3. AUC values, Sensitivity, Specivity, False Negativity and Positivity, NPV and PPV for the Ondokuz Mayis Nomogram and the 14 Other Models Validated in our Series, and the Number of Patients and AUC values by the Original Series

Models	No. of patients	AUC (95% CI) by the original series	AUC (95% CI) by our dataset	Sensitivity %	Specifity %	False positive rate %	False negative rate %	NPV %	PPV %
MSKCC	702	0.76*	0.840 (0.78-0.89)	80	79	22	19	81	78
Mayo	551	0.77 (0.72-0.81)	0.739 (0.67-0.80)	66	72	31	31	69	69
Louisville	1253	0.68 (0.64-0.71)	0.701 (0.63-0.76)	86	46	40	22	78	60
Cambridge	118	0.84 (0.75-0.86)	0.808 (0.75-0.86)	70	80	24	27	73	76
Stanford	213	0.85 (0.81-0.89)	0.819 (0.76-0.87)	83	74	25	18	82	75
SNUH (Seoul)	184	0.82 (0.76-0.88)	0.803 (0.74-0.86)	74	78	24	24	76	76
MOU (Masaryk)	330	0.76 (0.70-0.82)	0.836 (0.78-0.88)	79	70	21	22	78	79
Tenon	561	0.81 (0.79-0.83)	0.781 (0.72-0.83)	75	66	32	26	74	68
Turkish	607	0.80*	0.826 (0.77-0.88)	78	80	20	20	80	80
Ljubljana	460	0.79 *	0.812 (0.75-0.86)	78	75	26	22	78	74
Rome	139	0.77 (0.69-0.86)	0.754 (0.69-0.81)	76	70	30	25	75	70
MD Anderson	506	0.80 (0.75-0.84)	0.856 (0.80-0.90)	79	81	20	20	80	80
DEU (9 Eylul)	170	0.81*	0.844 (0.79-0.89)	84	75	25	17	83	75
European	1000	0.75 (0.66-0.76)	0.750 (0.68-0.81)	69	69	32	30	70	68
Ondokuz Mayis	237	0.87 (0.82-0.91)	0.871 (0.82-0.91)	85	78	22	15	85	78

* 95% confidence interval (CI) was not reported. NPV: negative predictive value, PPV: positive predictive value

Discussion

Our findings demonstrate that the non-SLNs were negative in 49% of patients with positive SLNs. If ALND had not been performed, half of the patients with positive SLNs would not have undergone ALND without any benefit for staging, outcome, or decision-making for adjuvant therapy. Furthermore, these patients would not have been exposed to the potential morbidity of ALND. The main problem is determining which patients with positive SLNs should undergo completion ALND and which should not. To find a solution to this problem, predictive factors for non-SLN metastasis among patients with positive SLN have been investigated, and many nomograms or scoring systems have been developed to predict the likelihood of non-SLN metastasis using predictive factors (Van Zee et al., 2003; Degnim et al., 2005; Chapgar et al., 2006; Pal et al., 2008; Kohrt et al., 2008; Cho et al., 2008; Coufal et al., 2009; Coutant et al., 2009; Gur et al., 2010; Perhavec et al., 2010; Lombardi et al., 2011; Mittendorf et al., 2012; Meretoja et al., 2012; Derici et al., 2012).

Our study revealed that pT, LVI, ENE, negSLN, SLNMS, and multifocality were independent predictive factors associated with positive non-SLNs. These factors were used to create the present Ondokuz Mayis model. Age, tumor size, LVI, ENE, number of removed SLNs, negSLN, posSLN, pRatio, SLN metastasis size, ER, PR, and c-erb B2 status, detection methods for SLN, and preoperative ultrasonography have previously been studied as predictive factors for positive non-SLN in the described models (Van Zee et al., 2003; Degnim et al., 2005; Chapgar et al., 2006; Pal et al., 2008; Kohrt et al., 2008; Cho et al., 2008; Coufal et al., 2009; Coutant et al., 2009; Gur et al., 2010; Perhavec et al., 2010; Lombardi et al., 2011; Mittendorf et al., 2012; Meretoja et al., 2012; Derici et al., 2012).

The present Ondokuz Mayis model has an excellent discrimination capacity, with an AUC of 0.87, and to the

best of our knowledge, this is the highest AUC among the models tested to date. An AUC of 0.50 indicates no discrimination, 0.70 to 0.80 indicates acceptable discrimination, and 0.81 to 0.90 indicates excellent discrimination (AUC \geq 0.90 is rare) (Hosmer and Lemeshow, 2000). Inspection of the calibration curve of our model and the Hosmer-Lemeshow test suggest that our model fits and was well-calibrated, and there were no significant difference between the predicted and the observed probabilities. The definition of a subgroup with a low predicted value in the Ondokuz Mayis model showed that the false negative rates with a predicted probability (cut-off value) of \leq 10% and \leq 20% were 0% and 9%, respectively. Therefore, 67 (28%) or 36 (15%) patients at the predicted probability of \leq 20% or \leq 10%, respectively, could have been spared the ALND in our series. The AUC and the calibration quantification are important for the evaluation and validation of the models. However, as suggested by Degnim et al. (Degnim et al., 2005) and Coutant et al. (Coutant et al., 2009), the number of patients and the false negative rate at the selected low predictive probability level (i.e., \leq 10%) should also be considered. Degnim et al. (Degnim et al., 2005) stated that although the MSKCC model yielded an excellent AUC of 0.86 with the Michigan dataset, the FN rate at the predicted probability of 5% or less was 14%.

The MSKCC model is the first described nomogram (Van Zee et al., 2003). Following the MSKCC model described by Van Zee et al., many models have been proposed and tested by many studies (Van Zee et al., 2003; Kocsis et al., 2004; Degnim et al., 2005; Smidt et al., 2005; Soni et al., 2005; Lambert et al., 2006; Alran et al., 2007; Ponzzone et al., 2007; Cho et al., 2008; Klar et al., 2008; Kohrt et al., 2008; Pal et al., 2008; Poirier et al., 2008; Coufal et al., 2009; Coutant et al., 2009; Coutant et al., 2009; Scow et al., 2009; Gur et al., 2010; Moghaddam et al., 2010; Lombardi et al., 2011; Hessman et al., 2011; Hidar et al., 2011; Tan et al., 2011; Chen et al., 2012; Derici et al., 2012; Zhu et al., 2012; Sasada et al., 2013). The

last proposed international model from multiple centers in Europe, however, has not been validated to date. To the best of our knowledge, the present study is the first to validate the Rome and European models (Lombardi et al., 2011; Meretoja et al., 2012). Applying the MD Anderson model to our dataset demonstrated that it was the most suitable model for our patients, generating an AUC of 0.86, which is the second highest AUC after the present model. Furthermore, the scatterplot of the predicted probabilities generated by the MD Anderson and our model displayed a high agreement. The false negative rates at the predicted probability of 10% or less and 20% or less were 0% for the MD Anderson model. However, the number of patients at these probabilities were as low as 4% and 15%, respectively.

The most unsuitable models for our dataset were the Mayo, Louisville, Rome, and European models, which yielded AUC values of 0.73, 0.70, 0.75, and 0.75, respectively. The scatter plots of the predicted probabilities of the Ondokuz Mayıs and these 4 models showed a low agreement with a wide scatter around the line of agreement. Looking at the scatterplots, Degnim et al. (Degnim et al., 2005) reported that the MSKCC and Mayo models did not exhibit high levels of agreement for individual patients based on the Mayo dataset. The MSKCC nomogram, which was validated by the present dataset, showed a high AUC of 0.84, and the scatter plots of the predicted probabilities generated by the Ondokuz Mayıs and MSKCC models displayed a modest agreement. The false negative rate at a predicted probability of 10% or less was 0% for the MSKCC nomogram, but the number of patients (4%) was very low at this level.

The selection of patients with a low predicted probability level of ≤ 20 with a false negative rate of $< 10\%$ could spare ALND for an important subset of patients with SLN metastasis. The mean predicted probability for patients without non-SLN metastasis in our model was 28%. Could the selection of the higher level of predicted probability increase the risk of local recurrence? Zakaria et al. (Zakaria et al., 2008) reported that a group of patients with SLN metastasis with a mean MSKCC predicted probability of non-SLN metastasis of 20% who did not undergo ALND were found to be without any axillary lymph node recurrences with a mean follow-up 30 months. Morrow et al. (Morrow et al., 2009) also stated that omitting ALND for patients with less than a 30% risk of additional nodal involvement would likely not lead to axillary recurrence resulting in breast cancer-related death.

The European model based on our dataset yielded a moderate AUC of 0.71, which is the same as in the original series, and the scatter plot showed that the agreement with our model was low. Moreover, when the European model was applied to our dataset, the false negative rate with a predicted probability of 10% or less was 28%. These findings demonstrate that the European model (EM) was not suitable for our patients. The EM includes SLN size as ITC, micro- and macrometastasis, and does not include the largest SLN metastasis size in mm. In addition, the EM included a factor called "prevalence of non-SLN metastases", which was not included in our model or in the other models.

To date, the MSKCC nomogram has been the most tested and validated model, and the range of AUCs in different datasets have been reported as 0.58 to 0.86 (Van Zee et al., 2003; Kocsis et al., 2004; Degnim et al., 2005; Soni et al., 2005; Smidt et al., 2005; Lambert et al., 2006; Alran et al., 2007; Ponzzone et al., 2007; Cho et al., 2008; Klar et al., 2008; Kohrt et al., 2008; Pal et al., 2008; Poirier et al., 2008; Coufal et al., 2009; Coutant et al., 2009; Scow et al., 2009; Gur et al., 2010; Moghaddam et al., 2010; Hessman et al., 2011; Hidar et al., 2011; Lombardi et al., 2011; Tan et al., 2011; Chen et al., 2012; Derici et al., 2012; Zhu et al., 2012; Sasada et al., 2013). Among the studies validating MSKCC model, our dataset generated the second highest AUC of 0.85, following the AUC of 0.86 yielded by the Michigan dataset. Kocsis et al. (Kocsis et al., 2004) reported that the MSKCC nomogram could not have been validated in their dataset. Pal et al. (Pal et al., 2008), Coufal et al. (Coufal et al., 2009), Moghaddam et al. (Moghaddam et al., 2010), Tan et al. (Tan et al., 2011), and Chen et al. (Chen et al., 2012) have also reported AUCs of < 0.70 , thus not validating the MSKCC nomogram. The Stanford nomogram, which was the second most validated nomogram with a range of AUCs of 0.65-0.76, has also been validated by our patient series with an AUC of 0.82, which is the highest AUC among the studies validating the Stanford nomogram. Of the 13 studies, 8 validated the Stanford nomogram. The present study was the first to validate the Rome and European nomograms with AUC values of 0.71 and 0.75, respectively.

All models were validated by our dataset with an AUC of > 0.70 . This means that all of the models are of clinical utility for our patients if AUC is considered the only criteria. The MSKCC, Stanford, MOU (Masaryk), Turkish, Ljubljana, MD Anderson, and DEU models exhibited AUCs of > 0.80 and had excellent discrimination in our dataset. Considering the excellent AUC values, the agreement of predictive probabilities by scatterplot, and the number of patients and low false-negative rates with a selected predictive probability, we concluded that the MD Anderson model is the most suitable and could be used for our patients. However, our Ondokuz Mayıs model yielded the highest AUC among all models. The calibration was good, and the model had 0% and 9% false negative rates at probabilities of ≤ 10 for 36 (15%) and $\leq 20\%$ for 67 (28%) patients in our patient series. The application of the present Ondokuz Mayıs model for our patients could spare ALND for 36 (15%) or 67 (28%) patients depending on the choice of predicted probability by the patient and the surgeon. Mittendorf et al. (2012) reported that only 4 (4%) of 101 patients with a predicted probability of 10% or less had a positive non-SLN in the MD Anderson model, whereas applying MD Anderson model to our dataset showed that there were only 10 (4%) patients with a 0% false negative rate for the same predicted probability. Barranger et al. (2008) reported that the chances of having negative non-SLNs were 97.3% in patients with a score of 3.5 or less (median score) in the Tenon model. The application of the Tenon scoring system in our series showed that the probability of having negative non-SLNs with a score of 3.5 or less was 92%. However, the median Tenon score was 6 in our series. There were 155 (65%) patients with

median scores of 6 or less in our series with the application of Tenon scoring, and the false negative rate was 34%. These findings demonstrate that the number of patients and the false negative rates for a predicted probability in the models are different for different datasets.

In the present study, the Ondokuz Mayıs nomogram median cut-off value (predicted probability) was 46%. For patients in our series, the chance of having negative non-SLNs with a median cut-off value of 46% or less was 81%. With a value of 20% or less, this chance was 91%, and with a 10% cut-off value, the chance was 100%. The number of patients in the 20% or less or the 10% or less categories in our series were 67 (28%) and 36 (15%) of patients, respectively. To make a decision about whether to avoid or perform ALND, along with the false negative rates (i.e., the chance of having positive non-SLNs), the chance of having negative non-SLNs (i.e., the false positive rate) for a given predicted probability should also be considered (Hidar et al., 2011).

The present nomogram is based on a relatively small sample size and has not been validated by another series. Almost all patients underwent simultaneous lumpectomy or mastectomy and frozen section SLN analysis in our clinic. Therefore, some predictive factors were not available before the SLN analysis, and a second surgery for ALND may be required if the present nomogram was used.

The present Ondokuz Mayıs model has an excellent discrimination capacity to distinguish patients at low risk for positive non-SLN from high risk patients, with an AUC of 0.87, and to the best of our knowledge, this is the highest AUC among the models tested to date. Our model also fits and was well-calibrated. Our findings showed that the false negative rates with a predicted probability of $\leq 10\%$ and $\leq 20\%$ were 0% and 9%, respectively. Therefore, 67 (28%) or 36 (15%) patients at the predicted probability of $\leq 20\%$ or $\leq 10\%$, respectively, could have been spared the ALND in our series.

Nomograms, which are methods to predict the possibility of non-SLN metastasis, do not yet have the ability to replace ALND. They are increasingly being used by many surgeons (Park et al., 2007). As Scow et al. (2009) noted, models always perform best in the population on which they are based. Thus, all nomograms may not have utility for all patient populations. Every clinic should validate a model before using it, or in the best case, every clinic should create a nomogram, analyze it and consider the characteristics of the nomogram in making decisions about omitting or performing ALND. The characteristics of the nomogram could also be shared with patients and their families during counseling before surgery. The present nomogram to calculate the likelihood of non-SLN metastasis for SLN-positive breast cancer patients is available at http://tip.omu.edu.tr/bc_nomogram.ondokuzmayis/

References

Alran S, De Rycke Y, Fourchette V, et al (2007). Validation and limitations of use of a breast cancer nomogram predicting the likelihood of nonsentinel node involvement after positive

sentinel node biopsy. *Ann Surg Oncol*, **14**, 2195-201.

- Barranger E, Morel O, Coutant C (2009). Axilla scoring systems predicting risk of non-sentinel-node metastasis in breast cancer patients with a positive sentinel node. *Ann Surg Oncol*, **15**, 1261-2.
- Chagpar AB, Scoggins CR, Martin RC, et al (2006). Prediction of sentinel lymph node-only disease in women with invasive breast cancer; University of Louisville Breast Sentinel Lymph Node Study. *Am J Surg*, **192**, 882-7.
- Chen K, Zhu L, Jia W, et al (2012). Validation and comparison of models to predict non-sentinel lymph node metastasis in breast cancer patients. *Cancer Sci*, **103**, 274-81.
- Cho J, Han W, Lee JW, et al (2008). A scoring system to predict nonsentinel lymph node status in breast cancer patients with metastatic sentinel lymph nodes: a comparison with other scoring systems. *Ann Surg Oncol*, **15**, 2278-86.
- Coufal O, Pavlík T, Fabian, et al (2009). Predicting non-sentinel lymph node status after positive sentinel biopsy in breast cancer: what model performs the best in a Czech population? *Pathol Oncol Res*, **15**, 733-40.
- Coutant C, Olivier C, Lambaudie E, et al (2009). Comparison of models to predict nonsentinel lymph node status in breast cancer patients with metastatic sentinel lymph nodes: a prospective multicenter study. *J Clin Oncol*, **27**, 2800-8.
- Coutant C, Rouzier R, Fondrinier E, et al (2009). Validation of the Tenon breast cancer score for predicting non-sentinel lymph node status in breast cancer patients with sentinel lymph node metastasis: a prospective multicenter study. *Breast Cancer Res Treat*, **113**, 537-43.
- Degnim AC, Reynolds C, Pantvaidya G, et al (2005). Nonsentinel node metastasis in breast cancer patients: assessment of an existing and a new predictive nomogram. *Am J Surg*, **190**, 543-50.
- Derici S, Sevinc A, Harmancioglu O, et al (2012). Validation of three breast cancer nomograms and a new formula for predicting non-sentinel lymph node status. *Asian Pac J Cancer Prev*, **13**, 6181-5.
- Giuliano AE, McCall L, Beitsch P, et al (2010). 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*, **252**, 426-32.
- Gur AS, Unal B, Ozbek U, et al (2010). Validation of breast cancer nomograms for predicting the nonsentinel lymph node metastases after a positive sentinel lymph node biopsy in a multi-center study. *Eur J Surg Oncol*, **36**, 30-5.
- Hanley JA, McNeil BJ (1982). The meaning and use of the area under a receiver operating characteristic (ROC) curve. *Radiology*, **143**, 29-36.
- Hessman CJ, Naik AM, Kearney NM, et al (2011). Comparative validation of online nomograms for predicting nonsentinel lymph node status in sentinel lymph node-positive breast cancer. *Arch Surg*, **146**, 1035-40.
- Hidar S, Harrabi I, Benregaya L, et al (2011). Validation of nomograms to predict the risk of nonsentinel lymph node metastases in North African Tunisian breast cancer patients with sentinel node involvement. *Breast*, **20**, 26-30.
- Hosmer DW, Lemeshow S (2000). Applied Logistic Regression. 2nd edition. New York, NY: John Wiley.
- Kell MR, Burke JP, Barry M, et al (2010). Outcome of axillary staging in early breast cancer: a metaanalysis. *Breast Cancer Res Treat*, **120**, 441-447.
- Klar M, Jochmann A, Foeldi M, et al (2008). The MSKCC nomogram for prediction the likelihood of non-sentinel node involvement in a German breast cancer population. *Breast Cancer Res Treat*, **112**, 523-31.

- Kleinbaum DG, Kupper LL, Muller KE, Nizam A (1998). Applied regression analysis and other multivariable methods. 3rd ed. Pacific Grove, CA: Duxbury Press.
- Kocsis L, Svébis M, Boross G, Sinkó M, et al (2004). Use and limitations of a nomogram predicting the likelihood of non-sentinel node involvement after a positive sentinel node biopsy in breast cancer patients. *Am Surg*, **70**, 1019-24.
- Kohrt HE, Olshen RA, Bermas HR, et al (2008). New models and online calculator for predicting non-sentinel lymph node status in sentinel lymph node positive breast cancer patients. *BMC Cancer*, **8**, 66.
- Lambert LA, Ayers GD, Hwang RF, et al (2006). Validation of a breast cancer nomogram for predicting nonsentinel lymph node metastases after a positive sentinel node biopsy. *Ann Surg Oncol*, **13**, 310-20.
- Lombardi A, Maggi S, Lo Russo M, et al (2011). Non-sentinel lymph node metastases in breast cancer patients with a positive sentinel lymph node: validation of five nomograms and development of a new predictive model. *Tumori*, **97**, 749-55.
- Meretoja TJ, Leidenius MH, Heikkilä PS, et al (2012). International multicenter tool to predict the risk of nonsentinel node metastases in breast cancer. *J Natl Cancer Inst*, **104**, 1888-96.
- Mittendorf EA, Hunt KK, Boughey JC, et al (2012). Incorporation of sentinel lymph node metastasis size into a nomogram predicting nonsentinel lymph node involvement in breast cancer patients with a positive sentinel lymph node. *Ann Surg*, **255**, 109-15.
- Moghaddam Y, Falzon M, Fulford L, Williams NR, Keshtgar MR (2010). Comparison of three mathematical models for predicting the risk of additional axillary nodal metastases after positive sentinel lymph node biopsy in early breast cancer. *Br J Surg*, **97**, 1646-52.
- Morrow M (2009). Patterns of care with a positive sentinel node: echoes of an opportunity missed. *Ann Surg Oncol*, **16**, 2429-30.
- Pal A, Provenzano E, Duffy SW, et al (2008). A model for predicting non-sentinel lymph node metastatic disease when the sentinel lymph node is positive. *Br J Surg*, **95**, 302-9.
- Park J, Fey JV, Naik AM, Borgen PI, Van Zee KJ, Cody HS 3rd (2007). A declining rate of completion axillary dissection in sentinel lymph node-positive breast cancer patients is associated with the use of a multivariate nomogram. *Ann Surg*, **245**, 462-8.
- Perhavec A, Perme MP, Hocevar M, Besić N, Zgajnar J (2010). Ljubljana nomograms for predicting the likelihood of non-sentinel lymph node metastases in breast cancer patients with a positive sentinel lymph node. *Breast Cancer Res Treat*, **119**, 357-66.
- Poirier E, Sideris L, Dubé P, Drolet P, Meterissian SH (2008). Analysis of clinical applicability of the breast cancer nomogram for positive sentinel lymph node: the canadian experience. *Ann Surg Oncol*, **15**, 2562-7.
- Ponzone R, Maggiorotto F, Mariani L, et al (2007). Comparison of two models for the prediction of nonsentinel node metastases in breast cancer. *Am J Surg*, **193**, 686-92.
- Sasada T, Kataoka T, Shigematsu H, et al (2013). Three models for predicting the risk of non-sentinel lymph node metastasis in Japanese breast cancer patients. *Breast Cancer*, [Epub ahead of print].
- Scow JS, Degnim AC, Hoskin TL, Reynolds C, Boughey JC (2009). Assessment of the performance of the Stanford Online Calculator for the prediction of nonsentinel lymph node metastasis in sentinel lymph node-positive breast cancer patients. *Cancer*, **115**, 4064-70.
- Smidt ML, Kuster DM, van der Wilt GJ, Thunnissen FB, Van Zee KJ, Strobbe LJ (2005). Can the Memorial Sloan-Kettering Cancer Center nomogram predict the likelihood of nonsentinel lymph node metastases in breast cancer patients in the Netherlands? *Ann Surg Oncol*, **12**, 1066-72.
- Soni NK, Carmalt HL, Gillett DJ, Spillane AJ (2005). Evaluation of a breast cancer nomogram for prediction of non-sentinel lymph node positivity. *Eur J Sur Oncol*, **31**, 958-64.
- Tan EY, Ho B, Chen JJ, et al (2011). Predictors of nonsentinel nodal involvement to aid intraoperative decision making in breast cancer patients with positive sentinel lymph nodes. *ISRN Oncol*, **2011**, 539503.
- Van Zee KJ, Manasseh DM, Bevilacqua JL, et al (2003). A nomogram for predicting the likelihood of additional nodal metastases in breast cancer patients with a positive sentinel node biopsy. *Ann Surg Oncol*, **10**, 1140-51.
- Zakaria S, Pantvaidya G, Reynolds CA, et al (2008). Sentinel node positive breast cancer patients who do not undergo axillary dissection: are they different? *Surgery*, **143**, 641-7.
- Zhu L, Jin L, Li S, et al (2013). Which nomogram is best for predicting non-sentinel lymph node metastasis in breast cancer patients? A meta-analysis. *Breast Cancer Res Treat*, **137**, 783-95.