Validation of the Skåne University Hospital nomogram for the preoperative prediction of a disease-free axilla in patients with breast cancer

S. Majid1,2, P.-O. Bendahl3, L. Huss1,4, J. Manjer1,2, L. Ryden2,5 and L. Dihge5,6,*

1Department of Clinical Sciences Malmö, Lund University, Malmö, Sweden
2Department of Surgery, Skåne University Hospital, Lund-Malmö, Sweden
3Department of Oncology and Pathology, Clinical Sciences, Lund University, Sweden
4Department of Surgery, Helsingborg Hospital, Helsingborg, Sweden
5Department of Clinical Sciences Lund, Lund University, Lund, Sweden
6 Department of Plastic and Reconstructive Surgery, Skåne University Hospital, Malmö, Sweden

*Correspondence to: Department of Clinical Sciences Lund, Surgery, Faculty of Medicine, Lund University, Medicon Village, SE-223 81, Lund, Sweden (e-mail: looket.dihge@med.lu.se)

Preliminary findings were published as part of S.M.’s doctoral thesis.

Abstract

Background: Axillary staging via sentinel lymph node biopsy (SLNB) is performed for clinically node-negative (N0) breast cancer patients. The Skåne University Hospital (SUS) nomogram was developed to assess the possibility of omitting SLNB for patients with a low risk of nodal metastasis. Area under the receiver operating characteristic curve (AUC) was 0.74. The aim was to validate the SUS nomogram using only routinely collected data from the Swedish National Quality Registry for Breast Cancer at two breast cancer centres during different time periods.

Method: This retrospective study included patients with primary breast cancer who were treated at centres in Lund and Malmö during 2008–2013. Clinicopathological predictors in the SUS nomogram were age, mode of detection, tumour size, multifocality, lymphovascular invasion and surrogate molecular subtype. Multiple imputation was used for missing data. Validation performance was assessed using AUC and calibration.

Results: The study included 2939 patients (1318 patients treated in Lund and 1621 treated in Malmö). Node-positive disease was detected in 1008 patients. The overall validation AUC was 0.74 (Lund cohort AUC: 0.75, Malmö cohort AUC: 0.73), and the calibration was satisfactory. Accepting a false-negative rate of 5 per cent for predicting N0, a possible SLNB reduction rate of 15 per cent was obtained in the overall cohort.

Conclusion: The SUS nomogram provided acceptable power for predicting a disease-free axilla in the validation cohort. This tool may assist surgeons in identifying and counselling patients with a low risk of nodal metastasis on the omission of SLNB staging.

Introduction

Assessment of patients’ axillary lymph node (ALN) status and the number of metastatic lymph nodes is essential for planning the treatment of primary breast cancer and underscores the importance of accurate nodal staging. Nodal staging via sentinel lymph node biopsy (SLNB) is routinely performed for all patients with a clinically negative axilla. The oncological safety of SLNB has been widely documented, even at a false-negative rate (FNR) of 5–10 per cent1–3. However, early detection via public mammography screening programmes has decreased the rate of node-positive disease4. Thus, invasive staging typically reveals a disease-free axilla (N0) in most primary breast cancer patients, and SLNB yields no therapeutic benefit. Moreover, although SLNB is a minimally invasive procedure and is associated with significantly less postoperative morbidity than axillary lymph node dissection (ALND), surgical axillary staging results in an incidence rate of 5–7 per cent for lymphoedema, 11 per cent for arm pain/numbness and 23 per cent for quality-of-life impairment5.

The benefit of extensive nodal staging is debatable, as advances in adjuvant therapy are tailored to tumour features rather than pathological nodal status6,7. The need for complete ALND when SLNB confirms limited disease has been questioned based on the lack of adverse effects on survival when ALND was omitted during randomized trials, such as the IBCSG 23–01 trial8 and the ACOSOG Z0011 trial9 (with a 10-year median follow-up period). There is also a growing interest in the reduction of axillary surgical staging and SLNB omission for selected low-risk patients based on the ongoing prospective European SOUND10 and INSEMA11 randomized trials. In these trials, breast cancer patients who are candidates for breast-conserving surgery with clinically and ultrasonographically negative axillae are randomized to...
Fig. 1 The Skåne University Hospital nomogram for predicting a disease-free axilla versus any nodal metastasis

The present study aimed to validate the SUS nomogram in a population-based cohort from breast cancer centres in Lund and Malmo, using only routinely collected data from The National Quality Registry for Breast Cancer (NKBC) of Sweden\(^2\). This study also aimed to determine the per cent reduction in unbeneficial SLNB that could be achieved using the SUS nomogram to identify patients with the lowest risk of nodal metastasis, who might be spared from surgical axillary staging.

### Methods

The retrospective study protocol was approved by the regional ethical review board (Lund, Sweden; Dnr 2013/821). All model validation procedures are reported in accordance with the EQUATOR guidelines for the transparent reporting of diagnostic studies, the TRIPOD statement\(^23\) and the STARD 2015 checklist\(^24\).

### Validation cohort

This validation study included patients with primary invasive breast cancer who underwent breast surgery and axillary staging at two breast cancer centres in Lund and Malmö (Skåne University Hospital) between January 2008 and December 2013. The validation cohort was stratified according to site (Lund and Malmö) and time period: overlapping the development period for the original SUS nomogram (2009–2012); 1 year before and 1 year after the nomogram development period (2008 and 2013); and the entire validation period (2008–2013), which corresponded to seven separate validation analyses. As in the original report regarding the development and internal validation of the SUS nomogram, the exclusion criteria were as follows: male sex, presence of bilateral tumours, history of invasive breast cancer or in situ ductal carcinoma, and neoadjuvant chemotherapy. Patients with missing information regarding ALN status or the procedure for surgical axillary nodal staging were also excluded. In accordance with the Swedish National Guidelines for Breast Cancer\(^25\) at the time, ALND was recommended if SLNB revealed micrometastasis or macrometastasis.

### Table

| Variable                        | Score | Predicted probability |
|---------------------------------|-------|-----------------------|
| Tumour size (mm)                | 0–3   | 0.01–0.3              |
| Detected by mammographic screening | 0–1   | 0.3–0.7               |
| Age (years)                     | 0–10  | 0.1–0.9               |
| Multifocality                   | 0–1   | 0.05–0.9              |
| Vascular invasion               | 0–1   | 0.01–0.9              |
| Subtype*                        | 1–5   | 0.01–0.9              |

**Fig. 1** The Skåne University Hospital nomogram for predicting a disease-free axilla versus any nodal metastasis. The total score for each patient is assigned by drawing a vertical line from the appropriate point for each predictor down to the score scale and summing these scores. The predicted probability of NO is determined by drawing a vertical line from the total score scale up to the predicted probability scale in the lower part of the nomogram. *Subtypes: 1, luminal A-like; 2, luminal B-like (LumB)/human epidermal growth factor receptor 2 (HER2)-negative; 3, LumB/HER2-positive; 4, HER2-positive/non-luminal; 5, triple-negative.
Data collection and definition of clinicopathological determinants

Data regarding the clinicopathological determinants for predicting a disease-free axilla (N0 versus N+) using the SUS nomogram were retrieved from the NKBC, which covers over 99 per cent of all new breast cancer cases in Sweden. The determinants included in the nomogram were as follows: age, mode of detection, tumour size, multifocality, vascular invasion and St Gallen surrogate molecular subtype (Fig. 1). In addition, data were collected regarding baseline variables, such as tumour histopathological grade, ALN status and the auxiliary staging procedure. Pathological predictors included Ki-67 status, HER2 status, oestrogen receptor (ER) status, progesterone receptor (PR) status, multifocality and vascular invasion, which were defined according to the Swedish Society of Pathology classification system. Positive nodes (N+) were defined as ALNs with macrometastasis or micrometastasis, and negative nodes (N0) were defined as nodes with isolated tumour cells or no metastasis. Vascular invasion was defined as the presence of tumour cell invasion through a vessel wall and endothelium or was defined based on the presence of tumour cells in vascular spaces or in the underlying endothelium of vascular channels. Multifocality was defined as the presence of two or more tumours in the same breast, which were at least 20 mm apart and separated by normal tissue or an in situ carcinoma.

The surrogate molecular subtype categories in the SUS nomogram were based on the proposed classification from the 13th St Gallen International Breast Cancer Conference in 2013, which defined five different subtypes based on the ER, PR, HER2 and Ki-67 statuses. The validation analyses defined hormone-receptor positive (LumB/HER2-: ER+, HER2-, Ki-67 low), luminal B-like/HER2 negative (LumB/HER2-: ER+, HER2-, Ki-67 high), luminal B-like/HER2 positive (Lumb/HER2+: ER+, HER2+, any Ki-67), HER2-positive/non-luminal (HER2+:non-luminal: ER-, HER2+, any Ki-67) and triple-negative (TNBC: ER-, HER2-, any Ki-67).

Statistical analysis

Multiple imputation was applied to handle missing data regarding the determinants for the SUS nomogram. The following variables were included in the imputation models:

- The dichotomous outcome (N0 versus N+)
- All determinants for the SUS nomogram, except for the surrogate molecular subtype (i.e., age, tumour size, mode of detection, vascular invasion and multifocality)
- The individual components of the surrogate molecular subtype (i.e., ER status, HER2 status and Ki-67 status)
- Other potential predictors of missingness or the missing values themselves (e.g., menopausal status, histological grade, PR status, treating centre, date of diagnosis and date of surgery).

The patterns of missing data were investigated and used in the specification of the imputation models. Two hundred new complete imputed data sets were generated, and 20 iterations were used to create each of the imputed datasets. Continuous variables were imputed using linear regression, and categorical variables were imputed using predictive mean matching. Classification into surrogate molecular subtypes was performed following imputation of the individual components.

The performance of the nomogram for discriminating between N0 and N+ was analysed using a receiver operating characteristics (ROC) analysis and by summarizing the AUC values averaged over the 200 imputed datasets. The value of the linear predictor (LP), which is the predicted log odds of lymph node negativity for patients in the validation cohort when applying the weights from the logistic regression model underlying the SUS nomogram, was used in the ROC analysis:

\[
LP = -1.793 + 0.166 \times (Lumb/HER2-) + 0.102 \times (Lumb/HER2+) + 0.389 \times (HER2 + /non - luminal) + 1.620 \times (TNBC) + 0.021 \times (age in years) + 0.562 \times (screening detected) - 0.059 \times (tumour size in mm) + 0.542 \times (no multifocality) + 1.542 \times (no vascular invasion)
\]

The first four variables in the equation are dummy variables comparing the subtype in the parentheses (coded 1) to the luminal A-like subtype as the reference subtype (coded 0). In addition, three binary predictors (screening detected disease, no multifocality and no vascular invasion) are coded 1 for logical yes and 0 for logical no.

The inverse logit transformation was subsequently used to calculate the predicted probability of N0:

\[
p = \exp(LP)/(1 + \exp(LP))
\]

A pooled AUC was calculated as the mean of the 200 imputation-specific AUC estimates. The variance of this estimate, which is the sum of the variances within and between imputations, was calculated using Rubin’s rule. The Hosmer–Lemeshow test and the corresponding calibration plot were applied to evaluate the goodness of fit of the logistic regression model underlying the SUS nomogram in the validation dataset. The predicted probabilities (p) of N0 were divided into deciles, and the observed proportions of patients with a disease-free axilla were compared to the average predicted probabilities in each decile. Thus, 10 dots on a line with a 45° slope would reflect perfect calibration. The calibration plots for the present study have 200 dots per decile, which correspond to one dot per imputed dataset.

The cut-off point for predicting a disease-free axilla was set at the maximized negative predictive value (NPV, the NPV closest to 100 per cent), which would identify patients with a very low probability of nodal metastasis who would not be likely to benefit from surgical axillary staging via SLNB. The mean values from the 200 imputed datasets, which were rounded to the nearest integer, for true-positive (TP), true-negative (TN), false-positive (FP), and false-negative (FN) results were calculated and used to calculate the FNR as FN/(FN + TP). The possible SLNB reduction rates were defined as (TN + FN)/(TN + FP + FN), based on a maximized NPV and maximum FNRs of 5 per cent and 10 per cent to reflect the acknowledged FNRs for SLNB staging. All calculations were performed using IBM SPSS® software (version 25.0; IBM Corp., Armonk, New York, USA) and Stata® software (version 16.0; StataCorp, College Station, Texas, USA).

Results

Between January 2008 and December 2013, 3979 patients were diagnosed with breast malignancy and scheduled for treatment at the SUS centres in Lund and Malmö. Patients were excluded if they were of male sex (30 patients), had a history of invasive breast cancer (43) or in situ carcinoma (126), were confirmed to
have received neoadjuvant treatment (170), any uncertainty regarding whether neoadjuvant therapy was administered (2), or missing information regarding the procedure for axillary nodal staging (189) or the final nodal status (316). Thus, the final study population included 2939 eligible patients: 1318 patients who were treated in Lund and 1621 patients who were treated in Malmö (Fig. 2). The clinicopathological characteristics and tumour detection modes for the overall validation cohort, the Lund centre and the Malmö centre are shown in Table 1.

Validation cohort characteristics
The axillary staging procedure involved SLNB for 1801 patients (61.3 per cent), nodal sampling for 18 patients (0.6 per cent) and ALND for 1115 patients (37.9 per cent). Axillary metastasis was detected in 1008 patients (34 per cent) in the overall validation cohort (Table 1).

Relative to the Lund validation cohort, the Malmö validation cohort had tumours that were more frequently detected based on symptomatic presentation (52 versus 42 per cent; P < 0.001) and that had high Ki-67 expression (54 versus 43 per cent; P = 0.023). Vascular invasion was more commonly observed in the Lund cohort, whereas the Malmö cohort had a higher proportion of HER2+ tumours.

Interestingly, both centres had high proportions of missing Ki-67 data; the Lund centre had a higher proportion of cases with missing vascular invasion data, and the Malmö centre had a higher proportion of cases with missing HER2 data. These differences might be explained by different registration routines at the two centres. Thus, a sensitivity analysis was performed by recoding the missing HER2 values as non-amplified to validate the predictive performance of the SUS nomogram. The HER2 sensitivity analysis did not substantially alter the nomogram’s predictive ability (data not shown).

External validation of the SUS nomogram
For predicting a disease-free axilla (N0 versus N+), the original SUS nomogram had an AUC value of 0.74 (95 per cent c.i. 0.70 to 0.79) based on detailed clinicopathological data that were extracted from the medical records of patients who were diagnosed between January 2009 and December 2012 in Lund (development cohort, 598 patients). Internal validation using bootstrap datasets revealed slight optimism in that AUC, based on a minor decline in the discriminative performance (–0.009).

Using only routinely collected registry data, the AUC value for the same period was 0.76 (95 per cent c.i. 0.72 to 0.80) in the Lund cohort and 0.73 (95 per cent c.i. 0.70 to 0.76) in the Malmö cohort (Table 2). The AUC value for the time period 1 year before and 1 year after the SUS nomogram development period (i.e., 2008 and 2013) was 0.75 (95 per cent c.i. 0.70 to 0.81) in the Lund cohort and 0.74 (95 per cent c.i. 0.70 to 0.79) in the Malmö cohort.
Table 1: Comparison of baseline patient and tumour characteristics in the validation cohort

|                             | All (Lund and Malmö) 2008–2013 | Lund 2008–2013 | Malmö 2008–2013 | P‡ | Lund versus Malmö |
|-----------------------------|---------------------------------|----------------|-----------------|----|--------------------|
| No. (row percentage)        | 2939 (1318 (45))                | 1621 (55)      |                 |    |                    |
| Age (years)*                 | 62 (24–96)                      | 62 (24–91)     | 62 (29–96)      | 0.664§<0.001 |
| Mode of detection            |                                 |                |                 |    |                    |
| Mammography screening        | 1518 (53)                       | 734 (58)       | 784 (48)        |    |                    |
| Symptomatic presentation     | 1360 (47)                       | 523 (42)       | 837 (52)        |    |                    |
| Missing                      | 61                              | 61             | 0               |    |                    |
| Tumour size (pT)†            |                                 |                |                 | 0.004¶|                    |
| ≤20 mm (pT1)                | 1877 (67)                       | 840 (70)       | 1037 (65)       |    |                    |
| >20–≤50 mm (pT2)             | 860 (31)                        | 338 (28)       | 522 (33)        |    |                    |
| >50 mm (pT3)                | 61 (2)                          | 22 (2)         | 39 (2)          |    |                    |
| Missing                      | 141                             | 118            | 23             |    |                    |
| Histological grade           |                                 |                |                 | 0.258¶|                    |
| I                           | 649 (22)                        | 281 (22)       | 368 (23)        |    |                    |
| II                          | 1258 (43)                       | 556 (43)       | 702 (44)        |    |                    |
| III                         | 993 (34)                        | 457 (35)       | 536 (33)        |    |                    |
| Missing                      | 39                              | 24             | 15             |    |                    |
| Multifocality                |                                 |                |                 | 0.956|                    |
| Absent                      | 1804 (79)                       | 679 (79)       | 1125 (79)       |    |                    |
| Present                     | 480 (21)                        | 180 (21)       | 300 (21)        |    |                    |
| Missing                      | 655                             | 459            | 196            |    |                    |
| Vascular invasion           |                                 |                |                 | <0.001|                    |
| Absent                      | 1565 (85)                       | 194 (76)       | 1371(86)        |    |                    |
| Present                     | 280 (15)                        | 62 (24)        | 218 (14)        |    |                    |
| Missing                      | 1094                            | 1062           | 32             |    |                    |
| Oestrogen receptor           |                                 |                |                 | 0.199|                    |
| Positive                    | 2444 (87)                       | 1062 (88)      | 1382 (86)       |    |                    |
| Negative                    | 366 (13)                        | 146 (12)       | 220 (14)        |    |                    |
| Missing                      | 129                             | 110            | 19             |    |                    |
| Progesterone receptor        |                                 |                |                 | 0.828|                    |
| Positive                    | 2088 (74)                       | 900 (75)       | 1188 (74)       |    |                    |
| Negative                    | 720 (26)                        | 307 (25)       | 413 (26)        |    |                    |
| Missing                      | 131                             | 111            | 20             |    |                    |
| HER2 status                  |                                 |                |                 | <0.001|                    |
| Not amplified                | 1587 (84)                       | 955 (88)       | 632 (78)        |    |                    |
| Amplified                   | 312 (16)                        | 129 (11)       | 183 (23)        |    |                    |
| Missing                      | 1040                            | 234            | 806            |    |                    |
| Ki-67                       |                                 |                |                 | 0.023|                    |
| ≤20%                        | 232 (50)                        | 93 (57)        | 139 (46)        |    |                    |
| >20%                        | 233 (50)                        | 70 (43)        | 163 (54)        |    |                    |
| Missing                      | 2474                            | 1155           | 1319           |    |                    |
| Regional lymph node metastases|                               |                |                 | 0.436|                    |
| N0                          | 1931 (66)                       | 856 (65)       | 1075 (66)       |    |                    |
| N+                          | 1008 (34)                       | 462 (35)       | 546 (34)        |    |                    |

Values in parentheses are column percentages unless indicated otherwise; * median (range). † According to the TNM classification for breast cancer. N0, disease-free axilla; N+, any lymph node metastasis. ‡ Pearson’s χ² test, except § Mann–Whitney U test; ¶ χ² test for trend.

Table 2: Performance of the Skåne University Hospital nomogram for prediction a disease-free axilla versus any nodal metastasis in the development cohort and the validation cohort

| Cohort                              | Time period | Data source | n    | AUC            |
|-------------------------------------|-------------|-------------|------|---------------|
| Development cohort                   |             |             |      |               |
| Lund, original SUS nomogram         | 2009–2012   | Medical records | 598  | 0.74 (0.70–0.79) |
| Validation cohort (all)              |             |             |      |               |
| Lund and Malmö                       | 2008–2013   | NKBC data   | 2939 | 0.74 (0.72–0.77) |
| Validation cohort (subgroups)        |             |             |      |               |
| a) Lund                              | 2009–2012   | NKBC data   | 916  | 0.76 (0.72–0.80) |
| b) Malmö                             | 2009–2012   | NKBC data   | 1086 | 0.73 (0.70–0.76) |
| c) Lund                              | 2008, 2013  | NKBC data   | 402  | 0.75 (0.70–0.81) |
| d) Malmö                             | 2008, 2013  | NKBC data   | 535  | 0.74 (0.70–0.79) |
| e) Lund                              | 2008–2013   | NKBC data   | 1318 | 0.75 (0.72–0.79) |
| f) Malmö                             | 2008–2013   | NKBC data   | 1621 | 0.73 (0.71–0.76) |

Values in parentheses are 95% confidence intervals. The validation cohort is stratified according to site (Lund and Malmö) and time period. a and b, overlapping the development period of the original Skåne University Hospital (SUS) nomogram (2009–2012); c and d, 3 year before and 1 year after the nomogram development period (2008 and 2013); e and f, the entire validation period (2008–2013). AUC, area under the receiver operating characteristic curve; NKBC, The National Quality Registry for Breast Cancer of Sweden.
The average AUC value for the entire validation period (i.e., 2008–2013) was 0.75 (95 per cent c.i. 0.72 to 0.79) in the Lund cohort and 0.73 (95 per cent c.i. 0.71 to 0.76) in the Malmo cohort.

Calibration plots stratified according to site (Lund and Malmo) and time period revealed satisfactory agreement between the predicted probability of N0 and the observed proportion of patients with a disease-free axilla, especially among cases with a high probability of N0 (Fig. 3). However, the model overestimated the risk of nodal metastasis in patients with a high probability of N+ (low probability of N0), based on the lower-left cluster of dots.
Implications of the SUS nomogram for reduction in SLNB

An NPV-orientated cut-off was assessed to evaluate the ability of the SUS nomogram to identify patients with a very low probability of nodal metastasis, who might not benefit from axillary staging via SLNB (Fig. 4). The mean maximized NPV was 95 per cent (FNR of 0.10 per cent) (Table 3). If alternative maximum FNRs of 5 and 10 per cent were used as cut-offs, which reflect the accepted FNRs for the SLNB procedure, the possible SLNB reduction rates were in the range of 15–26 per cent (Fig. 4 and Table 3).

Discussion

This study validated the ability of the SUS nomogram to predict a disease-free axilla (N0 versus N+) in a non-selected population-based cohort of 2939 consecutive cases. The discriminative ability in the validation cohort (AUC: overall, 0.74; the Lund centre, 0.75, and the Malmö centre, 0.73), which only included routinely collected registry data, was similar to that of the original SUS nomogram in the development cohort (internally bootstrap-validated AUC: 0.74)\(^2\). In this context, models’ predictive performances tend to decline in validation sets because of overfitting to the derivation data, which is known as optimism. Although several internal validation techniques address this issue, external validation is necessary to address the generalizability of a model’s predictive ability\(^2\). In general, an AUC value of 0.70–0.80 is thought to confirm that a model has good predictive power.\(^10\) Thus, the external validation findings (AUCs of 0.73–0.75), based on two centres with some differences in clinicopathological characteristics, indicate that the SUS nomogram is generalizable and may be applied to breast cancer patients from populations with similar characteristics. It may also be useful in populations with a different prevalence of lymph node negativity (64 per cent in the development cohort and 66 per cent in the validation cohort) and/or in cohorts with a different case mix. However, such applications would require recalibration of the predictor weights in the model.

A systematic review regarding the external validation of multivariable clinical prediction models concluded that calibration is an important measure of predictive performance that was missing in most studies\(^3\). This may at least partly explain the large number of published prediction models that have not been implemented in clinical practice. The present study revealed satisfactory calibration of the SUS nomogram, especially for patients with a low risk of nodal metastasis. Thus, the SUS nomogram may guide surgeons in counselling patients who are most likely

Table 3: Possible sentinel lymph node biopsy reduction rates within the validation cohort using the Skåne University Hospital nomogram for predicting a disease-free axilla at cut-offs corresponding to maximized negative predictive value and false-negative rates of 5 and 10 per cent to reflect the acceptable false-negative rates for sentinel lymph node biopsy

| Cut-off: Max NPV 0.95 | TP          | TN          | FP          | FN          |
|----------------------|-------------|-------------|-------------|-------------|
| No. SLNB Reduction Rate | (TN+FN)/(TP + TN + FP + FN) = 0.68% | 19          | 1912        | 1           |
| False-Negative Rate  | FN/(TP+FN) = 0.10%                     |             |             |             |

| Cut-off: NPV 0.89 | TP          | TN          | FP          | FN          |
|-------------------|-------------|-------------|-------------|-------------|
| No. SLNB Reduction Rate | (TN+FN)/(TP + TN + FP + FN) = 15.21% | 396         | 1535        | 51          |
| False-negative rate | FN/(TP+FN) = 5%     |             |             |             |

| Cut-off: NPV 0.87 | TP          | TN          | FP          | FN          |
|-------------------|-------------|-------------|-------------|-------------|
| No. SLNB Reduction Rate | (TN+FN)/(TP + TN + FP + FN) = 25.83% | 658         | 1273        | 101         |
| False-negative Rate | FN/(TP+FN) = 10%     |             |             |             |

Data on TP, TN, FP and FN are mean values from 200 imputed data sets rounded to the nearest integer. N0, disease-free axilla; N+, any lymph node metastasis; SLNB, sentinel lymph node biopsy; Max NPV, maximum negative predictive value; TP, True positive; TN, True negative; FP, False positive; FN, False negative.
to have a disease-free axilla and may be spared the SLNB procedure. However, the $P$ value from the Hosmer–Lemeshow test revealed fairly strong evidence against perfect calibration, indicating that the SUS nomogram seems to overestimate the risk of nodal spread when the risk of nodal metastasis is high, which, although not satisfactory, indicates appropriate nomogram cautiousness. Unsurprisingly, the best discrimination and calibration was observed for the Lund cohort that was treated during the same time period as the development cohort. Analyses in this subcohort validated the nomogram’s performance using only routinely collected registry data. The truly external validations were based on patients who were not treated at Lund during 2009–2012, which provide validation at the same site but during another time period as well as validation at another site during the same or different time periods. The truly external validation subsets provided generally similar results, with AUCs of 0.73–0.75 and calibration plots that exhibited the same pattern.

Although ALN status remains an important prognostic factor for primary breast cancer, the necessity of extensive surgical axillary staging has been questioned during the last decade. The IBCSG 23-0142 and ACOSOG Z001153 trials revealed that the omission of completion ALND was not inferior in terms of locoregional control or survival when the SLNB displayed limited metastasis in patients with a clinically node-negative axilla who underwent breast-conserving surgery, adjuvant breast irradiation and systemic treatment. Long-term follow-up of prospective randomized trials from the pre-SLNB era comparing ALND/axillary radiotherapy to no axillary surgery also revealed no significant difference in disease-free survival, although the rate of locoregional recurrence was somewhat increased when axillary surgery was not performed34–37.

The use of intensified adjuvant therapies over the last decade, which generally target tumour biological factors such as tumour molecular subtype, may further reduce the rates of axillary recurrence to 0.7–0.9 per cent after a benign SLNB result, as reported in early studies88,89. The present study evaluated whether the SUS nomogram could be used clinically to identify patients with a low risk of nodal metastasis who might benefit from the omission of surgical axillary staging, based on acknowledged FNR of 5–10 per cent for the SLNB technique. The results revealed possible SLNB reduction rates in the range of 15–26 per cent if FNR of 5–10 per cent were accepted for the prediction of disease-free axilla. Thus, the NPV-orientated cut-offs may be adjusted to guide surgeons in counselling patients regarding the omission of surgical axillary staging. The results are awaited of the randomized multicentric SOUND trial, which is comparing SLNB versus observation when the axillary ultrasound examination yields negative results6, and ultrasonographic axillary imaging with or without fine-needle biopsy remains a part of the routine preoperative work-up. Nevertheless, these imaging findings are influenced by the operator-dependent nature of axillary ultrasonography and the challenges of using imaging to determine accurately the nodal metastasis risk in low-burden disease. Therefore, the SUS nomogram may be an additional clinical tool for evaluating the nodal metastasis risk based on tumour biology, patient characteristics and the mode of detection. The SUS nomogram may provide additional information regarding axillary disease in this setting.

Since the publication of the Memorial Sloan-Kettering Cancer Center nomogram for estimating the risk of SLNB metastasis16, other nomograms have also been developed in more contemporary cohorts for predicting the risk of nodal metastasis18,40. However, the SUS nomogram is the first tool for predicting the likelihood of a disease-free axilla, for guiding the decision to omit axillary surgery for patients with the lowest likelihood of having nodal disease and to address potential SLNB reduction rates. Intrinsic tumour characteristics are known to be related to prognosis and locoregional control41,42, and the SUS nomogram incorporates the surrogate molecular subtypes to capture the lower risk of nodal metastasis for TNBC. Similarly, in the era of mammographic screening, the tumour detection mode also adds predictive value to the nomogram, alongside other key variables.

The most important limitation of the present study is the substantial proportion of missing data regarding some of the key pathological predictors in the validation cohort. This lack of completeness is a common problem in studies that use registry-based data sets. The predictors with the highest proportions of missing data were Ki-67 status and HER2 status, which are components of the surrogate molecular subtype, and vascular invasion. Multiple imputation was used to handle missing data, as this strategy provides unbiased estimates of missingness, and the missing values can thus be considered random, conditional on the other observed data, which is known as the missing at random assumption. This strategy has been shown to be effective even at high proportions of missing data43.

The highest true validation value (AUC: 0.75) suggests that the SUS nomogram does not perfectly predict a disease-free axilla, which highlights the complexity of lymphatic spread. Although more complex prediction models may have certain advantages for estimating nodal involvement44,45, the SUS nomogram is a readily available and user-friendly predictive tool in clinical settings. Studies based on retrospective registry data may be considered unreliable, given the risks of incomplete or improperly recorded data. Although the preoperative applicability of the original SUS nomogram depends on information obtained from histopathology results of core needle biopsy specimens, these validation results yielded the same AUC value. Thus, the nomogram’s utility was confirmed even with the use of only routine registry data with large amounts of missing data. The findings of this study are also strengthened by the use of a non-selected population-based validation cohort and the attempt to distinguish a disease-free axilla from any lymph node disease. The discrimination between node-negative cases and cases with any metastatic burden in the lymph nodes, including micrometastatic deposits, provides a more rigorous and cautious approach to risk estimation. Therefore, the SUS nomogram may help guide surgeons in counselling patients who are the least likely to have nodal metastasis and who may be safely spared from SLNB for axillary staging. Nevertheless, these results must be validated in other regions and patient populations to confirm that the SUS nomogram is universally applicable.

Funding

The study was supported by grants from the Skåne County Councils Research and Development Foundation, the Governmental Funding of Clinical Research within the National Health Service (ALF), the Swedish Cancer Society, the Erling Persson Family Foundation, the Gunnar Nilsson Cancer Foundation, the Einar and Inga Nilsson Foundation, Malmö University Hospital Cancer Research Fund, Skåne University Hospital Funds and Donations and the Central Hospital of Kristianstad.

The funders played no role in the study design, data collection and analysis, decision to publish, or preparation of the manuscript.
Acknowledgements

The authors thank the patients and clinicians who participated in The National Quality Registry for Breast Cancer.

Registration: This validation study was not preregistered in an independent institutional registry but is part of the NLS trial (Non-Invasive Lymph node Staging, registered 23 November 2018 with the ISRCTN registry: ISRCTN14341750).

Conflict of interest: The authors declare no conflicts of interest.

References

1. Bergkvist L, Frisell J, on behalf of the Swedish Breast Cancer Group the Swedish Society of Breast Surgeons. Multicentre validation study of sentinel node biopsy for staging in breast cancer. Br J Surg 2005; 92: 1221–1224
2. Lyman GH, Giuliano AE, Sorensen MR, Benson AB III, Bodurka DC, Burstein HJ et al. American Society of Clinical Oncology guideline recommendations for sentinel lymph node biopsy in early-stage breast cancer. J Clin Oncol 2005; 23: 7703–7720
3. Krag DN, Anderson SJ, Julian TB, Brown AM, Harlow SP, Ashikaga T et al.; National Surgical Adjuvant Breast and Bowel Project. Technical outcomes of sentinel-lymph-node resection and conventional axillary-lymph-node dissection in patients with clinically node-negative breast cancer: results from the NSABP B-32 randomised phase III trial. Lancet Oncol 2007; 8: 881–888
4. Lehtimaki T, Lundin M, Linder N, Sihto H, Holli K, Turpeenniemi-Hujanen T et al. Long-term prognosis of breast cancer detected by mammography screening or other methods. Breast Cancer Res 2013; 15: R134
5. Rao R, Euhus D, Mayo HG, Balch C. Axillary node interventions in breast cancer: a systematic review. JAMA 2013; 310: 1385–1394
6. Cardoso F, Kyriakides S, Ohno S, Penault-Llorca F, Poortmans P, Rubio IT et al.; ESMO Guidelines Committee. Early breast cancer: ESMO Clinical Practice Guidelines for diagnosis, treatment and follow-up. Ann Oncol 2019; 30: 3674
7. Kropp J, Ismaila N, Andre F, Bast RC, Barlow W, Collyar DE et al. Use of biomarkers to guide decisions on adjuvant systemic therapy for women with early-stage invasive breast cancer. American Society of Clinical Oncology Clinical Practice Guideline Focused Update. J Clin Oncol 2017; 35: 2838–2847
8. Gentilucci O, Veronesi U. Abandoning sentinel lymph node biopsy in early breast cancer? A new trial in progress at the European Institute of Oncology of Milan (SOUND: Sentinel node vs Observation after axillary UltraSouND). Breast 2012; 21: 678–681
9. Reimer T, Stachs A, Nekljudova V, Loibl S, Hartmann S, Wolter K et al. Restricted axillary staging in clinically and sonographically node-negative early invasive breast cancer (c/T1-2) in the context of breast conserving therapy: first results following commencement of the Intergroup-Sentinel-Mamma (INSEMA) Trial. Geburtshilfe Frauenheilkd 2017; 77: 149–157
10. Majid S, Tengrup I, Manjer J. Clinical assessment of axillary lymph nodes and tumor size in breast cancer compared with histopathological examination: a population-based analysis of 2,537 women. World J Surg 2013; 37: 67–71
11. Feng Y, Huang R, He Y, Lu A, Fan Z, Fan T et al. Efficacy of physical examination, ultrasound, and ultrasound combined with fine-needle aspiration for axilla staging of primary breast cancer. Breast Cancer Res Treat 2015; 149: 761–765
12. Dhige L, Grabau DA, Rasmussen RW, Bendahl PO, Ryden L. The accuracy of preoperative axillary nodal staging in primary breast cancer by ultrasound is modified by nodal metastatic load and tumor biology. Acta Oncol 2016; 55: 976–982
13. van Wely BJ, de Wilt JH, Francissens C, Teerenstra S, Strobbe LJ. Meta-analysis of ultrasound-guided biopsy of suspicious axillary lymph nodes in the selection of patients with extensive axillary tumour burden in breast cancer. Br J Surg 2015; 102: 159–168
14. Diepstraten SC, Sever AR, Buckens CF, Veldhuis WR, van Dalen T, van den Bosch MA et al. Value of preoperative ultrasound-guided axillary lymph node biopsy for preventing completion axillary lymph node dissection in breast cancer: a systematic review and meta-analysis. Ann Surg Oncol 2014; 21: 51–59
15. Lowes S, Leaver A, Cox K, Satchithananda K, Cosgrove D, Lim A. Evolving imaging techniques for staging axillary lymph nodes in breast cancer. Clin Radiol 2018; 73: 389–409
16. Bevilacqua JI, Kattan MW, Fey JV, Cody HS, 3rd, Borgen PI, Van Zee KJ. Doctor, what are my chances of having a positive sentinel node? A validated nomogram for risk estimation. J Clin Oncol 2007; 25: 3670–3679
17. van la Parra RF, Francissen CM, Peer PG, Ernst MF, de Roos WK, Van Zee KJ et al. Assessment of the Memorial Sloan-Kettering Cancer Center nomogram to predict sentinel lymph node metastases in a Dutch breast cancer population. Eur J Cancer 2013; 49: 564–571
18. Chen JY, Chen JJ, Yang BL, Liu ZB, Huang XY, Liu GY et al. Predicting sentinel lymph node metastasis in a Chinese breast cancer population: assessment of an existing nomogram and a new predictive nomogram. Breast Cancer Res Treat 2012; 135: 839–848
19. Holm-Rasmussen EV, Jensen MB, Balslev E, Kroman N, Tvedakov TF. Reduced risk of axillary lymphatic spread in triple-negative breast cancer. Breast Cancer Res Treat 2015; 149: 229–236
20. Liu N, Yang ZG, Liu XZ, Niu Y. Lymph node status in different molecular subtype of breast cancer: triple negative tumours are more likely lymph node negative. Oncotarget 2017; 8: 55334–5543
21. Dhige L, Bendahl PO, Ryden L. Nomograms for preoperative prediction of axillary nodal status in breast cancer. Br J Surg 2017; 104: 1494–1505
22. Regionala cancercentrum i samverkan. National Quality Registry for Breast Cancer (NKBC). https://www.cancercentrum.se/samverkan/cancerdiagnoser/brost/kvalitetsregister/ (accessed 2008–2013)
23. Collins GS, Reitsma JB, Altman DG, Moons KGM. Transparent Reporting of a multivariable prediction model for Individual Prognosis Or Diagnosis (TRIPOD): the TRIPOD Statement. Br J Surg 2015; 102: 148–158
24. Bossuyt PM, Reitsma JB, Bruuns DE, Gatsonis CA, Glasziou PP, Irwig L et al.; STARD Group. STARD 2015: an updated list of essential items for reporting diagnostic accuracy studies. Bmj 2015; 351: h5527
25. Regionala cancercentrum i samverkan. Swedish National Guidelines for Breast Cancer. https://www.cancercentrum.se/samverkan/cancerdiagnoser/brost/vardprogram/ (accessed 2008–2013)
26. Lofgren L, Eloranta S, Krawiec K, Asterkvist A, Lonnqvist C, Sandelin K; on behalf of the steering group of the National Register for Breast Cancer. Validation of data quality in the Swedish National Register for Breast Cancer. BMC Public Health 2019; 19: 495
