



OPEN ACCESS

EDITED BY

L. J. Muhammad,
Federal University Kashere, Nigeria

REVIEWED BY

Abbas Khan,
Shanghai Jiao Tong University, China
Khurshid Ahmad,
Yeungnam University, Republic of Korea

*CORRESPONDENCE

Looket Dihge
✉ looket.dihge@med.lu.se
Lisa Rydén
✉ lisa.ryden@med.lu.se

†These authors share last authorship

SPECIALTY SECTION

This article was submitted to
Breast Cancer,
a section of the journal
Frontiers in Oncology

RECEIVED 18 November 2022

ACCEPTED 13 February 2023

PUBLISHED 01 March 2023

CITATION

Dihge L, Bendahl P-O, Skarping I,
Hjærtström M, Ohlsson M and Rydén L
(2023) The implementation of NILS: A
web-based artificial neural network
decision support tool for noninvasive
lymph node staging in breast cancer.
Front. Oncol. 13:1102254.
doi: 10.3389/fonc.2023.1102254

COPYRIGHT

© 2023 Dihge, Bendahl, Skarping,
Hjærtström, Ohlsson and Rydén. This is an
open-access article distributed under the
terms of the [Creative Commons Attribution
License \(CC BY\)](https://creativecommons.org/licenses/by/4.0/). The use, distribution or
reproduction in other forums is permitted,
provided the original author(s) and the
copyright owner(s) are credited and that
the original publication in this journal is
cited, in accordance with accepted
academic practice. No use, distribution or
reproduction is permitted which does not
comply with these terms.

The implementation of NILS: A web-based artificial neural network decision support tool for noninvasive lymph node staging in breast cancer

Looket Dihge^{1,2*}, Pär-Ola Bendahl³, Ida Skarping^{3,4},
Malin Hjærtström³, Mattias Ohlsson^{5†} and Lisa Rydén^{1,6*†}

¹Department of Clinical Sciences Lund, Division of Surgery, Lund University, Lund, Sweden,

²Department of Plastic and Reconstructive Surgery, Skåne University Hospital, Malmö, Sweden,

³Department of Clinical Sciences Lund, Division of Oncology, Lund University, Lund, Sweden,

⁴Department of Clinical Physiology and Nuclear Medicine, Skåne University Hospital, Lund, Sweden,

⁵Department of Astronomy and Theoretical Physics, Division of Computational Biology and Biological Physics, Lund University, Lund, Sweden, ⁶Department of Surgery, Skåne University Hospital, Malmö, Sweden

Objective: To implement artificial neural network (ANN) algorithms for noninvasive lymph node staging (NILS) to a decision support tool and facilitate the option to omit surgical axillary staging in breast cancer patients with low-risk of nodal metastasis.

Methods: The NILS tool is a further development of an ANN prototype for the prediction of nodal status. Training and internal validation of the original algorithm included 15 clinical and tumor-related variables from a consecutive cohort of 800 breast cancer cases. The updated NILS tool included 10 top-ranked input variables from the original prototype. A workflow with four ANN pathways was additionally developed to allow different combinations of missing preoperative input values. Predictive performances were assessed by area under the receiver operating characteristics curves (AUC) and sensitivity/specificity values at defined cut-points. Clinical utility was presented by estimating possible sentinel lymph node biopsy (SLNB) reduction rates. The principles of user-centered design were applied to develop an interactive web-interface to predict the patient's probability of healthy lymph nodes. A technical validation of the interface was performed using data from 100 test patients selected to cover all combinations of missing histopathological input values.

Results: ANN algorithms for the prediction of nodal status have been implemented into the web-based NILS tool for personalized, noninvasive nodal staging in breast cancer. The estimated probability of healthy lymph nodes using the interface showed a complete concordance with estimations from the reference algorithm except in two cases that had been wrongly included (ineligible for the technical validation). NILS predictive performance to distinguish node-negative from node-positive disease, also with missing values, displayed AUC ranged from 0.718 (95% CI, 0.687-0.748) to 0.735 (95% CI, 0.704-0.764), with good calibration. Sensitivity 90% and specificity 34% were

demonstrated. The potential to abstain from axillary surgery was observed in 26% of patients using the NILS tool, acknowledging a false negative rate of 10%, which is clinically accepted for the standard SLNB technique.

Conclusions: The implementation of NILS into a web-interface are expected to provide the health care with decision support and facilitate preoperative identification of patients who could be good candidates to avoid unnecessary surgical axillary staging.

KEYWORDS

breast cancer, artificial intelligence, neural networks, disease diagnosis, lymphatic metastasis, sentinel lymph node, clinical decision support, programmable calculator

1 Introduction

Breast cancer in females is the most commonly diagnosed cancer, with approximately 2.3 million new annual cases worldwide (1). Breast cancer care has evolved into a multidisciplinary approach, with increasingly conservative surgical techniques in the breast and axilla and targeted systemic therapies (2). Evaluation of axillary nodal status in primary breast cancer determines the pathological stage of the disease, and the presence of axillary lymph node metastasis remains a key prognostic factor in breast cancer (3). Therefore, an accurate assessment of nodal status is still crucial in guiding the extent of axillary surgery and the need for neoadjuvant and adjuvant therapy including postoperative radiotherapy. Over the decades, several landmark trials have established a more conservative and less extensive approach for axillary nodal staging, from axillary nodal dissection (ALND) to today's standard staging by sentinel lymph node biopsy (SLNB) (4–6). Furthermore, previous routine completion of axillary lymph node excision in sentinel lymph node-positive breast cancer has been abandoned by the results of randomized trials on axillary management (7–10) and further advances in adjuvant treatment. While paradigm-changing studies such as the ACOSOG Z0011 study (10) indicate an increasing interest in limiting axillary surgery, and early detection by mammography screening has shifted the distribution of breast cancer toward earlier stages (11), SLNB is still routinely performed in patients with clinically node-negative breast cancer. The vast majority of these patients present with benign axillary lymph node status, and the invasive surgical procedure has no therapeutic benefit but increases the risk of morbidity related to pain, seroma, neuropathy, arm lymphedema, and postoperative infection (12–16). In 2021, ASCO published an updated guideline on axillary staging which included recommendations to forego SLNB in selected patients and endorsed a case-by-case evaluation to warrant patient-centered axillary treatment (17).

Although imaging methods for assessing the lymph nodes of the axilla have shown promising results (18–20), the diagnostic accuracy of imaging techniques alone has been insufficient to replace SLNB for nodal staging. Therefore, alternative methods for surgical staging of axillary lymph nodes have been presented, for example, prediction models that combine clinical and tumor

biological and/or radiological properties to estimate the probability of axillary lymph node metastasis. Machine learning-assisted models, including artificial neural networks (ANN), have been proposed to supplement standard statistical methods in classification tasks and decision support in cancer (21). An ANN model can handle non-predefined relations, such as nonlinear interactions, and does not require distributional assumptions. However, this technique comes at the cost of reduced interpretability and difficulty in gaining insight into the model features (22).

In 2019, our research group presented an ANN prototype for preoperative noninvasive lymph node staging (23). This original model utilized fifteen patient-related and clinicopathological characteristics to predict healthy lymph nodes and showed better discriminatory ability than a multivariable logistic regression model. The fifteen variables in the model were selected based on association to nodal status in previously published models, variable selection in the ANN prototype and their availability in routine breast cancer work-up. The potential to abstain from surgical axillary staging by SLNB was shown in 18–27% of newly diagnosed breast cancer patients. A health-economic decision-analytic model demonstrated that the implementation of the ANN prototype is associated with substantial cost reductions and overall health gains (24).

This study aims (1) to present the implementation of a further improved ANN algorithm for nodal status prediction in early breast cancer into a web-based tool for noninvasive lymph node staging (NILS), (2) to report on the predictive performance of the NILS tool with integrated ANN models to handle different access to preoperative data, (3) to assess the clinical utility by estimating possible SLNB reduction rates applying the NILS tool to predict nodal status in early breast cancer.

2 Materials and methods

2.1 Data collection for model development

The current NILS tool is a further development of an ANN prototype for the prediction of healthy lymph nodes published in

2019 (23). The study population for training and internal validation of the initial ANN algorithm was based on a prospectively maintained pathological database. We identified consecutive clinically node negative patients with invasive breast cancer who underwent primary breast surgery and surgical axillary lymph node staging between January 1, 2009, and December 31, 2012, at Skåne University Hospital, Lund, Sweden. Ethical permission was obtained from the local ethics committee at Lund University (LU 2012/340), and patients were allowed to opt out of the study. Exclusion criteria were: male gender, prior ipsilateral surgery in the breast or axilla, palpable axillary lymphadenopathy, neoadjuvant therapy, and unknown pathological nodal status after routine surgical axillary staging. Fifteen potential clinicopathological predictors of nodal metastases were identified. Information on the mode of breast cancer detection (detected at screening mammography vs. symptomatic presentation) was obtained from the mammography screening records. Histological classification and definition of axillary nodal status has previously been described (23). The following data were collected from medical records and surgical pathology reports: age at diagnosis, menopausal status, body mass index, bilateral disease, tumor localization in the breast, tumor multifocality, tumor size, histological type, histological grade, status of estrogen receptor (ER), progesterone receptor (PR), human epidermal growth factor receptor 2 (HER2), Ki67 value, and presence of vascular invasion (VI).

2.2 The ANN structure

2.2.1 Variable selection

The initial ANN algorithm for predicting healthy lymph nodes vs. metastatic lymph nodes (N0 vs. N+) previously reported in 2019 required multiple preoperative and postoperative clinicopathological inputs variables (23). Missing data were handled using multiple random imputations. The selection of variables from this first predictive ANN model provided a set of candidate predictors for the current NILS tool. The key features of the variable selection step are summarized here. Ranking of the input variables was performed by determining the input weights of the candidate predictors, and variable redundancy reduction was obtained by maintaining only the top-ranked variables for the prediction of healthy lymph nodes. The overall significance of the input variables was evaluated using a permutation technique. More specifically, a given input variable was decoupled from the patient by a random permutation across the evaluation cohort and the impact of this randomization on predictive ability was assessed. The predictor variable that is associated with the greatest reduction in predictive ability was assigned an importance value of one. All other predictors were given positions according to the associated decrease in the predictive performance at randomization. The 10 highest-ranking variables provided the contribution profile of the final input variables to the current NILS tool. The current tool included in order of importance: largest tumor size, VI, tumor multifocality, ER status, histological type, PR status, mode of detection, age at diagnosis, tumor localization, and Ki67 value.

2.2.2 Multilayer perceptron model

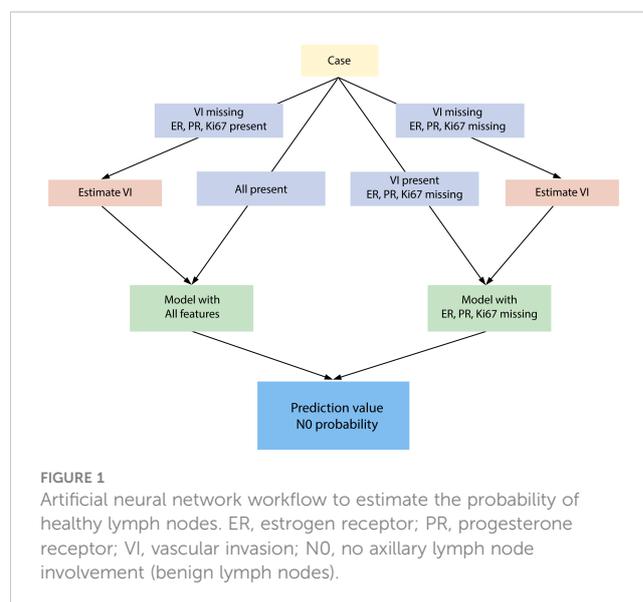
Ensemble models, each comprising 30 multilayer perceptron (MLP) neural networks, were used for the development and internal validation of predictive ANN models. An overview of the original ANN structure has been presented in a previous publication (23). For completeness, we present the model training and internal validation steps in detail.

Each MLP in an ensemble had one hidden layer with 10 nodes and a single output node encoding the probability of healthy axillary lymph nodes. The number of input nodes was 12–16, depending on the specific combination of input features. The activation function for the hidden nodes is a hyperbolic tangent function and a logistic function for the output node. An ensemble prediction was performed by averaging the individual MLP predictions.

2.2.3 Workflow for N0 prediction

The present NILS ANN algorithm was trained to handle combinations of missing data for 10 input variables. Strategies for handling missing input variables have been developed to reflect the clinical pathway of diagnostic workup in breast cancer, with a special focus on the risk of incomplete preoperative histopathological reporting. Therefore, the prediction model required that all variables should be present, except for VI and/or ER status, PR status, and Ki67 value. Figure 1 shows the N0 prediction workflow, which allows combinations of variable inputs.

In the first step, the application programming interface (API) verifies whether information on VI is missing. In cases with missing VI, one of the two imputation models was chosen to acquire an imputed value for VI. The choice of the VI imputation model was based on the presence of ER status, PR status, and Ki67 value. If any of these three input variables were missing, a second model without these inputs was applied. In the next step of N0 prediction after the inclusion of VI, one of the two models was selected in the same manner as described above. If any of the three input variables (ER status, PR status, and Ki67 value) were missing, a model without



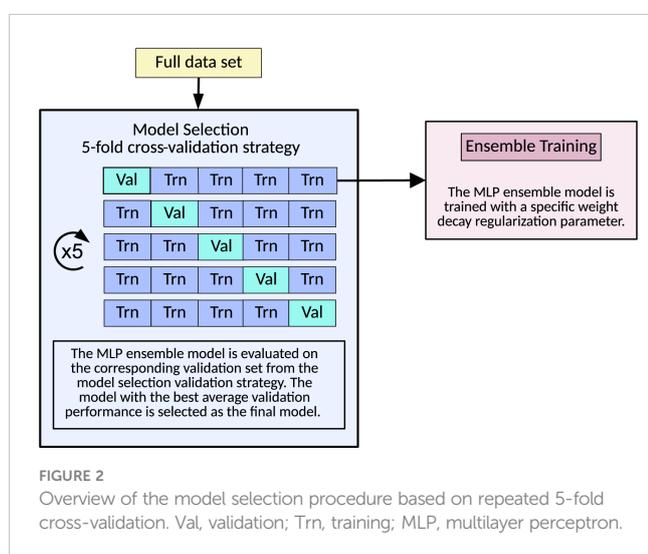
these features was applied; otherwise, an algorithm using all values was assigned.

2.2.4 Model training

For each model presented in the N0 prediction workflow (Figure 1), an ensemble of MLPs was trained. Each MLP was trained on a unique subsample of the original training dataset. Such a sample was created by first randomly splitting the original training dataset into three parts of equal size stratified by class. One part was removed from the data, leaving two parts that formed a training dataset. This procedure was repeated 10 times with different random splits into three parts. In total, 10×3 training datasets were created and used to train 30 MLPs in an ensemble.

2.2.5 Model selection

Each individual MLP was regularized using the weight decay method (25). This method uses a hyperparameter that sets the amount of regularization to be used. The same weight decay parameter was used for all MLPs in the ensemble, and the optimal parameters were determined using a model selection procedure. Model selection was accomplished using the method of repeated 5-fold cross-validation. The original training dataset was split randomly into five parts of equal size, with a preserved class fraction within each part. Each part was used to validate an ensemble model trained on the remaining four parts. This procedure was repeated five times, resulting in 5×5 validation results. The average of these 5×5 results was used to validate the performance of the ensemble model. This repeated 5-fold cross-validation procedure was performed for a sequence of different weight decay parameters. The weight decay parameter corresponding to the best 5-fold cross-validation performance was selected as the optimal parameter and was used in the final ensemble model. An overview of the model selection process is shown in Figure 2.



2.2.6 Data preprocessing

A maximum of 10 input variables contributed to 16 values used as inputs for the MLP according to the following criteria

- 1) Tumor size (mm): single numerical value. Normalized to unit standard deviation and zero mean before training.
- 2) VI: a single binary value. Present = 1, Absent = 0
- 3) Multifocality: Single binary value. Present = 1, Absent = 0
- 4) ER status: Single binary value. Positive = 1, Negative = 0
- 5) Histological type: categorical, three values. Invasive ductal cancer of no special type (NST) (1,0,0), invasive lobular cancer (ILC) (0,1,0), and histological subtypes other than pure NST or ILC (0,0,1).
- 6) PR status: Single binary value. Positive = 1, Negative = 0
- 7) Mode of detection: Single binary value. Mammographic screening = 1; symptomatic presentation = 0
- 8) Age (years): Single numerical value. Normalized to unit standard deviation and zero mean before training.
- 9) Tumor localization: semi-categorical, five values. Breast tumor localization was computed using two features, the right/left specification and the o'clock position of the tumor, and assigned as central or within the breast quadrants: upper inner (UIQ), upper outer (UOP), lower inner (LIQ), and lower outer (LOQ).
- 10) Ki67: Single numerical value. Normalized to unit standard deviation and zero mean before training.

2.3 Reporting on model predictive performance, calibration, and statistical analysis

The development of the NILS predictive tool for healthy lymph nodes and the report of findings were following an EQUATOR Guideline for reporting machine learning predictive models (26). The predictive performance of the model was represented by the area under the receiver operating characteristics curves (AUC). Sensitivity and specificity values at defined cut-off point were also presented. The reported performances of the different ANN models were obtained from the internal 5-fold procedure. No separate test set or additional cross-validation loop was used for more unbiased estimation of prediction performance. To further simplify the reporting of the performance measures, a single validation result was computed from the internal 5×5 fold. Each of the five validation lists was fused to a single validation list, incorporating all training data. Repeated cross-validation led to five validation lists, and a single final list was computed by averaging the predictions for each patient. The reported performance measures (e.g. AUC) were computed from the final validation list and the Hosmer-Lemeshow (HL) chi-squared statistic was used to assess the calibration. Predictive and calibration performances were evaluated using different ANN models within the current NILS tool.

The true positive (TP), true negative (TN), false positive (FP), and false negative (FN) outcomes were evaluated. The false negative rate (FNR) was computed as the number of FN cases divided by the number of cases with confirmed lymph node metastasis $FNR = FN / (FN + TP)$. Possible SLNB reduction rate, defined as $(TN + FN) / (TN + FN + TP + FP)$, was evaluated at a cut-off point equivalent to the maximum FNR of 10%, to reflect the FNR associated with the SLNB technique. Sensitivity $TP / (TP + FN)$ and specificity $TN / (TN + FP)$ values were assessed at the same cut-off point.

The distribution of clinicopathological variables and mode of detection across the nodal status outputs (comparison between N0 and N+) was assessed using Pearson's χ^2 test for categorical variables and Mann–Whitney U test for variables measured on a continuous scale. Custom-made software written in C (GCC) version 7.5.0, Perl version 5.26.1, SPSS Statistics for Windows version 25.0 (IBM, Armonk, New York, USA), and Stata version 17 (StataCorp, College Station, Texas, USA) were used for statistical calculations and graphics. A *P*-value should not be interpreted in relation to a cut-off point for significance, but as a continuous measure of evidence against the corresponding null hypothesis.

2.4 The benchmark criteria for N0 status reflect the accuracy of the SLNB technique

A cut-off point reflecting the accepted FNR of the SLNB procedure was selected to predict N0. This level of cut-off point, corresponding to 10% FNR, was intended to discriminate N0 vs. N+ and identify individuals with a low probability of axillary metastatic disease, where the omission of surgical axillary staging by SLNB would be supported by the NILS tool. Two cut-off point values were applied: one for the model with all input variables present and one for the model with missing input variables (ER status, PR status, and Ki67 value).

2.5 Development of a user-friendly interface for personalized prediction of N0 status

The principles of user-centered design were applied to develop an interactive web interface and provide a decision support tool to predict the patient's probability of healthy axillary lymph nodes. Specifically, the guiding principles of risk communication (27, 28) have been implemented in the web-based format of the NILS tool and involve the following:

- I) Presenting the numeric estimate of the probability of N0 for an individual patient.
- II) The estimate was interpreted in the context of the background distribution of the estimated probabilities of healthy lymph nodes in the breast cancer population used for model development.
- III) Conveying uncertainties and disclaimers related to current research status.

IV) Combining numerics with a visual display to aid comprehension.

V) Tailoring estimates with a personalized message for each individual output.

The design was developed in collaboration with Medos AB, an external medical software company (<https://www.medos.se>). Graphical features allow users to set up the NILS interface on computers and common mobile devices, such as smartphones and tablets.

2.6 Technical validation of the NILS interface

To validate that the implementation of the NILS tool followed the design specifications, the NILS interface (version 0.1.0) was validated with 100 reference test subjects, selected to cover different combinations of missing histopathological input data. The reference dataset was based on the development cohort, in which the following random patients were selected:

- 25 patients with a complete dataset.
- 25 patients with a complete dataset, excluding VI.
- 25 patients in whom the ER status, PR status, or Ki67 value is missing.
- 25 patients where VI is missing in conjunction with a missing value in at least one of ER status, PR status, or Ki67 value.

Clinicopathological data were assigned for manual input into the interface. The NILS tool calculated a prediction value and identified a cut-point for each datapoint in the reference dataset. The exact value from the NILS tool was obtained from web browser console output, since the value displayed in the graphical user interface is rounded to integer value. The results of the predictions made through NILS was compared to the results obtained through the reference ANN algorithms.

3 Results

3.1 Clinical and histological characteristics

Data from a consecutive cohort of 800 invasive breast cancer cases, one observation per case, were included in the training and internal validation of the original ANN prototype for healthy lymph nodes. The benign axillary nodal status was displayed in 516 (64.5%) cases. Patients with N0 breast cancer had smaller tumors and a higher rate of mammography-detected disease than those with N+ status (Supplemental Table 1). Furthermore, breast tumors of patients with N0 displayed a higher proportion of ER negative status, had lower levels of Ki67, and were more often lacking multifocality and VI.

3.2 Model performance

3.2.1 Predictive performance and calibration for the estimation of VI

In cases with missing data on VI, one of two separate ANN models was chosen to acquire an imputed value for VI. The model for VI imputation comprising complete input data (9 variables) achieved an AUC of 0.795 (95% confidence interval [CI], 0.751–0.835), [Figure 3](#). If any of the input variables, ER status, PR status, or Ki67 value were missing, an ANN model without these inputs was applied for VI prediction and displayed a discriminatory performance of AUC 0.788 (95% CI, 0.744–0.831). The corresponding median 8-degree of freedom (df) HL statistic values were 7.5 ($P = 0.480$) and 13.4 ($P = 0.100$), respectively.

3.2.2 Predictive performance and calibration for the prediction of N0

The present NILS prediction tool to distinguish healthy axillary lymph nodes with complete data and, if necessary, the prediction of VI displayed good discrimination with a calculated AUC of 0.735 (95% CI, 0.704–0.764). The NILS tool was well calibrated, with an observed 8-df-HL statistic of 9.9 ($P = 0.270$). The prediction model without input values for ER status, PR status, and Ki67 achieved a discriminatory performance of AUC 0.718 (95% CI, 0.687–0.748). The corresponding median 8-df-HL statistic value for this model was 9.7 ($P = 0.290$). [Figure 4](#) shows the discriminative ability (N0 vs. N+) of different ANN pathways within the NILS tool.

3.2.3 Performance of the NILS tool compared to the previous ANN prototype for N0

The initial ANN prototype presented in 2019 to distinguish healthy lymph nodes (N0 vs. N+), comprising 15 potential clinicopathological predictors, achieved a mean validation AUC of 0.735. The corresponding internally validated AUC for N0 was 0.740 (95% CI, 0.723–0.758) (23). Compared with the present NILS tool, which includes a reduced number of input variables and allows missing values for VI, ER, PR, and Ki67, there was a 0.022-point estimate difference between the AUCs, as shown in [Figure 4](#).

3.3 Implications for sentinel lymph node reduction rate

[Table 1](#) shows the possible SLNB reduction rates after applying the NILS tool to assign healthy lymph nodes (N0 vs. N+) and accepting missing values for VI, ER, PR, and Ki67. A SLNB reduction rate of 26% would be attained by applying the present NILS tool at a cut-off point corresponding FNR of 10% to discriminate N0. At this cut-off point, sensitivity and specificity values of 90% and 34%, respectively, were demonstrated.

3.4 The NILS interactive decision support tool

3.4.1 Background distribution curves

The workflow for the prediction of N0 probability consisted of the ANN model with complete data for all input variables, including a possible imputation of VI, and the ANN model with missing values for any of ER, PR, and Ki67. For each model, the output was the predicted probability (p) of healthy lymph nodes (N0). The smoothed empirical distribution of the predicted probabilities obtained during model development is shown in [Figure 5](#). The smoothing procedure includes a transformation of the empirical probabilities to logits, that is, $\ln(p/(1-p))$, where the predicted probability p is a value between 0 and 1, and thereafter kernel smoothing using the command `k-density` in the statistics package `Stata/MP 17.0`. A Gaussian kernel with a bandwidth of 0.5 was used. The logit transformation ensures that the support of the distribution remains within the interval $0 < p < 1$.

3.4.2 The interactive web-based NILS interface

The ANN algorithms for nodal status prediction have been implemented into a fully operative web-based decision support tool. The NILS interactive prediction tool for healthy lymph nodes was built to incorporate features of a user-centered design. Using 10 entry values with information on the patient's age, data from mammography examination, and core needle biopsies from the breast tumor, the tool can predict the likelihood of having healthy

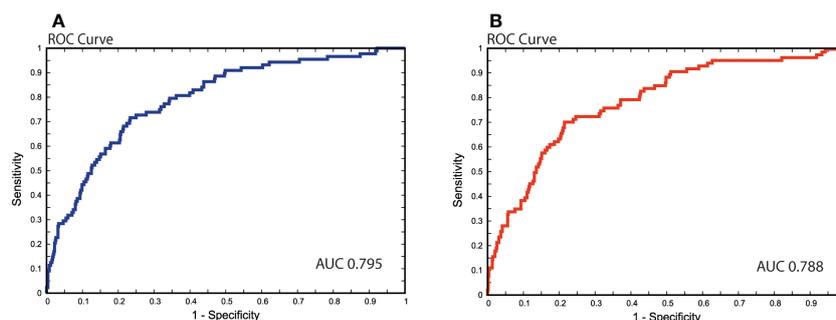


FIGURE 3

Receiver operating characteristics (ROC) curves visualizing predictive performance for estimating vascular invasion. Area under the curve (AUC) (A) Model comprising of complete input data (9 variables); (B) Model without the three input variables: estrogen receptor (ER), progesterone receptor (PR), and Ki67.

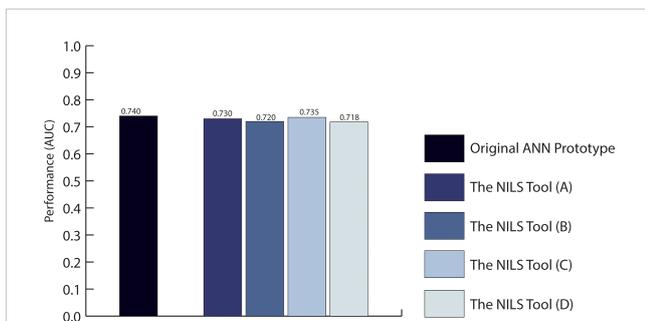


FIGURE 4 Predictive performance to discriminate healthy lymph node status at the time of breast cancer diagnosis. The diagram displays the area under the receiver operating characteristics curves (AUCs) for the original ANN prototype presented in 2019, which required 15 input variables and the ANN pathways within the present NILS tool, which requires the maximum number of 10 top-ranked input variables. (A) The NILS tool that includes all top-ranked variables and random imputation of missing data; (B) The NILS tool with top-ranked variables excluding estrogen receptor (ER), progesterone receptor (PR), and random imputation of missing data; (C) The NILS tool that includes all top-ranked variables, vascular invasion (VI) is estimated if missing, random imputation of the remaining missing data; (D) The NILS tool with top-ranked variables excluding ER, PR, and Ki67. VI is estimated if missing random imputation of the remaining missing data.

lymph nodes at the time of breast cancer diagnosis. The NILS interface includes clickable information icons that display explanations of the input fields, as shown in Figure 6. The ANN algorithms of the NILS tool predict the probability of healthy lymph nodes, which are displayed graphically together with the applicable cut-off point. Therefore, the outputs are presented numerically, as the patient’s individual estimated probability of healthy lymph nodes and graphically in relation to the set cut-off point for the binary outcome (N0 vs. N+). Individual estimation is visualized graphically as a vertical line superimposed on a smoothed histogram, which describes the distribution of estimated probabilities for healthy lymph nodes in the population used to develop the NILS tool. In addition, the suggested cut-off point is superimposed on this distribution. Figure 7 shows an output display that incorporates these features. The output score was compared to the specified cut-off point for binary classification (N0 vs. N+). The result of this comparison is positive if the prediction is above the cut-off point (the green icon assigns a probability of N0 above the cut-off point). If the prediction is below the cut-off point, the result is negative (the red icon assigns a probability of N0 below the cut-off point).

The NILS tool was submitted to an initial usability test to assess the clinical functionality of the interface and to identify necessary alterations to the layout and interactive elements, such as input and output fields, information icons, and explanatory information text. Each individual estimate of nodal status prediction includes a disclaimer to communicate the uncertainty related to the current research status. The NILS website has been translated into English and can be accessed at <https://nils.thep.lu.se/>.

To ensure quality assurance for future clinical preoperative applicability, the NILS predictive tool is now applied in a clinical trial for validation purposes without interfering with the current breast cancer workflow (ISRCTN registry, study ID 14341750) (29). The prediction tool or “calculator” is currently accessible within the research program by entering login credentials.

3.4.3 User interface validation

NILS version 0.1.0 was tested with a total of 100 reference test patients selected to cover combinations of missing histopathological input data. Two subjects with no information on multifocality were mistakenly included (did not fulfilled the eligibility criteria for the validation) and generated incorrect test results. For the rest, 98 tested subjects, the estimated probability of healthy lymph nodes made using the NILS tool matched the predictions obtained through the reference algorithm to the 8th decimal. The cut-point and graphical curve presented on the output screen were correct for all test subjects.

4 Discussion

In this study, we present an implementation of ANN algorithms for the prediction of healthy lymph nodes as a decision support tool for a personalized, noninvasive lymph node staging in breast cancer. The NILS tool is the first interactive tool based on ANN to complement the updated ASCO guideline, which recommend abstaining SLNB in selected patients on a case-by-case basis (17). The NILS tool allowed missing histopathological input values, and achieved predictive performance with an AUC ranging from 0.718 to 0.735 with good calibration. By applying the NILS tool for nodal status and setting the cut-off point for model classification to reflect the accepted FNR of the SLNB procedure, the potential to reduce surgical axillary intervention by 26% was demonstrated.

The clinical utility of the NILS tool for noninvasive prediction of healthy lymph nodes depends on the information obtained in the

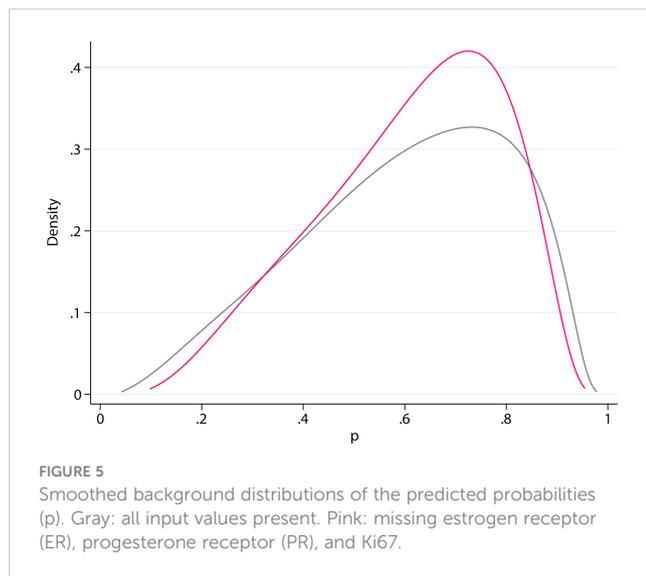
TABLE 1 Possible sentinel lymph node biopsy (SLNB) reduction rates at cut-point equivalent to the maximum false negative (FNR) rate of 10%, reflecting the accepted FNR of the SLNB procedure.

Model (N0 versus N+)	TP	TN	FP	FN	Sensitivity	Specificity	FNR	SLNB reduction rate
ANN Prototype*	258	190	324	28	91%	37%	<10%	27%
The NILS Tool	256	177	339	28	90%	34%	<10%	26%

N0, No axillary lymph node involvement (benign lymph nodes); N+, axillary lymph node positive (metastatic lymph nodes); TP, true positive; TN, true negative; FP, false positive; FN, false negative; FNR, false negative rate; SLNB, sentinel lymph node biopsy.

*Dihge L, Ohlsson M, Eden P, Bendahl PO, Ryden L. Artificial Neural Network Models to Predict Nodal Status in Clinically Node-Negative Breast Cancer. BMC cancer 2019 (23).

Comparison between potential SLNB reduction rates by applying the former ANN prototype (15 input variables)* and the current NILS tool (maximum 10 input variables) trained to tolerate missing preoperative input values for vascular invasion, estrogen receptor (ER), progesterone receptor (PR), Ki67.



preoperative setting. To handle the risk of incomplete preoperative histological reporting based on core needle biopsy (CNB), the ANN algorithm in the NILS tool incorporated strategies to allow missing data for VI and/or ER status, PR status, and Ki67 value. Previous studies have shown high specificity but low sensitivity for VI detection in CNB (30) with a reported 30% detection failure rate (31). Therefore, a separate model for estimating VI is integrated into the present NILS tool. CNB samples have been shown to be accurate in determining the ER and PR status (32), however, a significant difference between Ki67 values in CNB and surgical specimens has been observed, with a relative decrease from the biopsy to surgical sample (33). As high Ki67 expression has been associated with axillary nodal metastasis (34), the NILS tool may overestimate the risk of axillary nodal spread in the preoperative setting, indicating appropriate cautiousness of the tool.

Previously, models have been developed to predict sentinel lymph node status using a variety of clinical and pathological factors (35, 36). One of the first was the Memorial Sloan-Kettering Cancer Center nomogram based on logistic regression analysis to estimate the probability of SLNB metastasis, which provided an AUC of 0.754. The final variable selection in the presented NILS tool focused on clinical knowledge, previous literature and variables associated with nodal status in the present cohort. The ten highest-ranking variables were included: tumor size, vascular invasion, multifocality, ER status, histological type, PR status, mode of detection, age, tumor localization, and Ki67 all of which have repeatedly been reported to be of value for predicting nodal status in breast cancer. Machine learning-based algorithms for predicting axillary nodal spread have been proposed to reflect the complex multifactorial process of lymph node metastasis (37, 38). Especially, artificial intelligence with a deep learning approach in radiological image recognition, has gained attention to predict axillary nodal metastasis (39, 40). However, the clinical significance of these findings in routine preoperative settings remains limited. ANNs are usually included in the group of machine learning techniques described as black-box models and are often criticized as nontransparent. For a given output, it is typically difficult to gain a comprehensive understanding of the relative importance of each

input variable (41). Visualization techniques based on sensitivity analysis have been proposed (42). Consequently, to better comprehend the significance of each input variable in the ANN algorithm within the NILS tool, sensitivity analysis was previously applied to the variable selection process. In the publication from 2019, mean odds ratios was presented with the corresponding percentiles to highlight the dynamic nature of the 10 highest-ranking input variables for the prediction of nodal status (23).

Digital clinical decision support systems in breast cancer surgical care remains sparse (43). Diverse predictive methods has been applied on the benefit of adjuvant chemotherapy in the different systems, such as OncoType DX[®] embedded in a decision support algorithm (44) and cumulative hazard function by PREDICT (45). Hitherto, diagnostic decision support tools have not had as much impact as other types of clinical decision support systems (46). Reasons identified included negative perceptions in health care, inferior accuracy due to missing input data and poor integration of the tools with medical records (46, 47). Although the overall performance of the NILS tool was moderate (AUC 0.720–0.735), the tool displayed a high sensitivity (90%) to distinguish breast cancer patients most likely to harbor healthy lymph nodes. The ANN algorithms also allow different cut-off points to be assigned, which can be determined by the actual clinical setting. The reduction from the mandatory 15 input variables previously required (23) to 10 variables in the updated NILS tool, with an allowance of missing input entries with only 0.005–0.022 decrease in AUC point estimate, strengthens the clinical utility and reinforces the potential to spare one in four patients the surgical SLNB procedure.

Given the increasing evidence around de-escalating axillary surgery, the Choosing Wisely statement (48) declared that SLNB is not required for women ≥ 70 years with clinically node-negative (cN0) breast cancer, which is hormone receptor-positive and HER2-negative if they are adjuvantly treated with hormonal therapy. Other advocates omitted SLNB in all low-risk patients aged ≥ 75 years, even without planned hormonal treatment (49). Randomized trials, such as SOUND (50) and INSEMA (51), now address the possibility of omitting SLNB and include patients of all ages with ultrasonographically disease-free axillae. Although advances are promising, the sensitivity of routine axillary ultrasound alone is considered insufficient to replace SLNB (52). The updated ASCO guideline on the management of the lymph nodes in early breast cancer state that patients should be evaluated on a case-by-case basis to ensure patient-centered care (17). Consequently, decision support tools that consider the patient's individual preoperative clinicopathological variables for nodal status prediction could contribute to a better differentiation of breast cancer patients with disease-free axilla from those with nodal metastasis.

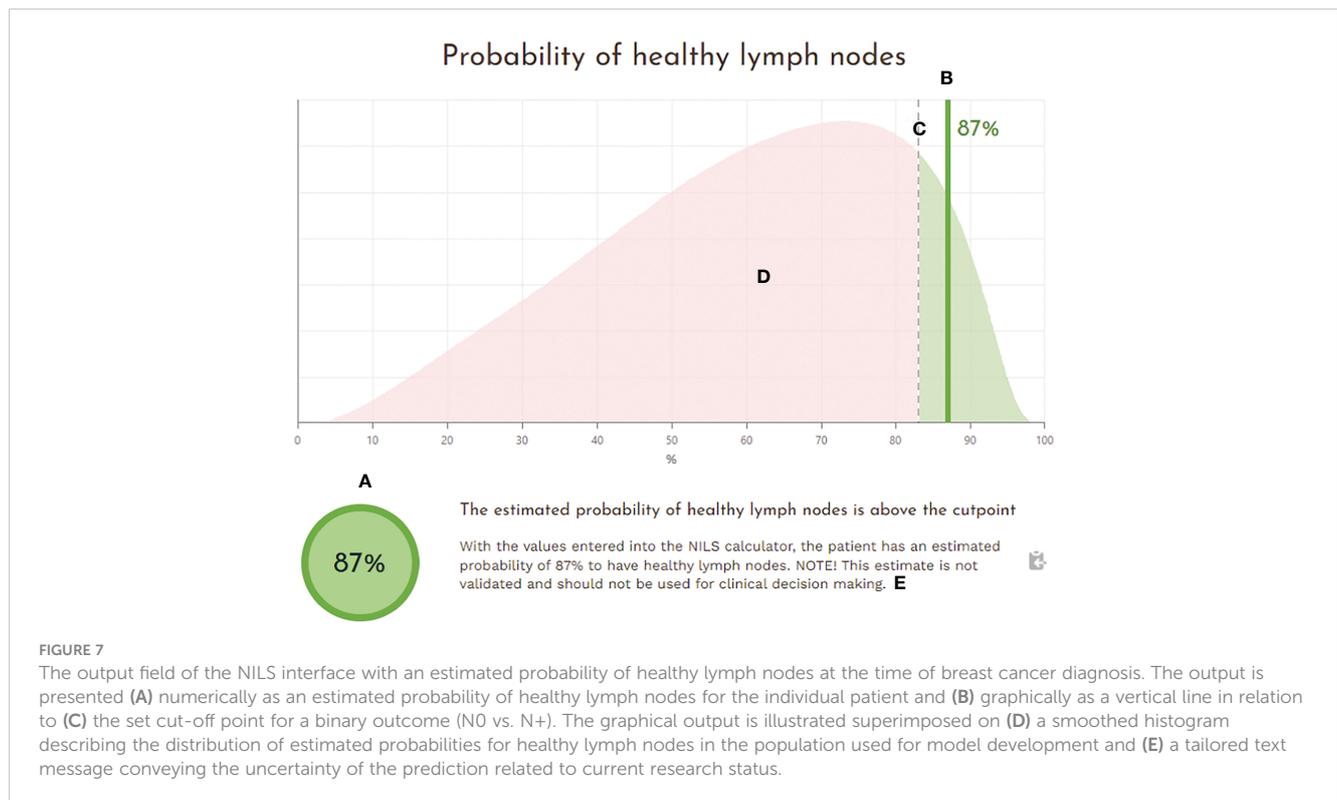
An important step before clinical implementation of a decision-support tool is to consider both associated health care costs and impact on patients' quality of life under plausible scenarios. Compared to the standard of care with SLNB, the adoption of the NILS tool for noninvasive staging of nodal status has been shown to promote reduced health care costs and gains in quality of life, especially in patients undergoing breast-conserving surgery (24).

FIGURE 6
The input screen of the online NILS interface. Clinicopathological and mammographical input entries for the prediction of healthy lymph nodes with essential user interface elements: (A) Clickable information icons with explanatory information text; (B) Clarifying illustration of the required input entry; (C) Option to choose “Unknown” or “Undefined” icons for input variables: vascular invasion, estrogen receptor (ER), progesterone receptor (PR), and Ki67.

Although the NILS prediction tool has been developed for clinical use following the framework of the Tripod checklist (53) to increase the methodological consistency and quality of the prediction models, the true validity of the model should be assessed in a fully independent dataset. To address this, a validation study is currently being conducted based on two external cohorts that assess the geographic and temporal validity of the NILS prediction tool (ISRCTN 14341750) (29).

Limitations of the NILS tool include that the model was developed using data mostly from patients with screen-detected breast cancer (57%, Supplemental Table 1). This is because Sweden has a population-based mammography screening program with an overall attendance of >80% (54). The NILS tool is not designed for settings with limited access to population-based screening, but the current tool can provide a range of probabilities of healthy lymph

nodes in patients with screen-detected *versus* symptom-detected breast cancers. Detection by mammography screening appears to be an independent factor associated with lower risk of axillary lymph node involvement (55). Another limitation is that variation in the proportions of node-positive breast cancer may also produce errors in the accuracy of the tool. However, procedures to adjust outputs of a classifier to new *a priori* probability has been proposed to mitigate these errors and might be applied in the future (56). A third limitation is that no differentiation was made between micrometastases and macrometastases using the current definition of nodal metastatic lesions (N+). While this may demonstrate the appropriate cautiousness of the NILS tool by appointing the classification task to differentiate healthy lymph node status from those harboring any metastasis, the finding of only micrometastatic deposits in sentinel nodes has no longer relevance



as a criterion for completion of axillary surgery (8, 57). Similarly, the decision on adjuvant systemic therapy is mainly based on tumor biomarkers and/or gene profiling rather than on identifying low-burden nodal metastasis (58). However, the presence of nodal macrometastasis in breast cancer remains relevant for decision-making regarding adjuvant radiotherapy (59). Ongoing research efforts to overcome this limitation include validating the NILS tool in datasets that specify the distinction between micrometastasis and macrometastasis in sentinel lymph nodes. Improving efficacy of systemic treatment regimens in breast cancer can pave the way to further expand de-escalation of axillary surgery and safe omission of SLNB for more patients. Thus, the NILS tool needs to be flexible to new target populations and adapt accordingly. The validity of the NILS tool will be assessed in patients undergoing different breast surgical approaches (breast-conserving surgery vs. mastectomy). Furthermore, the predictive performance of the NILS tool with additional features will be evaluated, including the incorporation of mammography-based radiomics model for the prediction of axillary lymph node status.

5 Conclusion

This study reports the implementation of ANN algorithms for axillary nodal status prediction as a web-based decision supporting tool for noninvasive lymph node staging. The NILS tool uses 10 routinely available entry values and allows missing histopathological data on core needle biopsies to predict the likelihood of benign lymph

nodes at the time of breast cancer diagnosis. The NILS tool is a personalized decision support tool based on ANN to complement the updated ASCO guideline, which recommends refraining from SLNB in selected low-risk patients after evaluation on a case-by-case basis. By applying the NILS tool for preoperative prediction of lymph node status, the potential to avoid surgical axillary staging was demonstrated in about one in four patients. In addition, we have shown substantial cost reductions and overall health gains associated with the implementation of the NILS tool. This study highlights the potential of ANN-based prediction tools in clinical use to aid in diagnosis and reduce unnecessary axillary surgery in early breast cancer.

Data availability statement

The data analyzed in this study is subject to the following licenses/restrictions: The datasets generated and/or analyzed during the current study are not publicly available because of privacy and ethical restrictions but are available from the corresponding authors (LD or LR) on reasonable request. Requests to access these datasets should be directed to LD, looket.dihge@med.lu.se or LR, lisa.ryden@med.lu.se.

Ethics statement

This study was approved by the Ethics Committee of Lund University (LU 2012/340). The study received ethical approval for

the use of an opt-out methodology and written informed consent for participation was not required for this non-interventional study in accordance with the national legislation and the institutional requirements.

Author contributions

Conceptualization, LR, MO, and LD; methodology, LR, LD, MO, and P-OB.; software, MO; validation, MO; formal analysis, MO, LD, and P-OB; investigation, LD, LR; resources, LR and MO; data curation, LD, MO, and P-OB; writing – original draft preparation, LD; writing – review and editing, MO, LR, P-OB, IS, and MH; visualization, LD, LR, and P-OB; supervision, LR and MO; project administration, LR; funding acquisition, LR. All authors contributed to the article and approved the submitted version.

Funding

This study received funding from The Swedish Breast Cancer Funds 2017/3502073; The Erling Persson Family Foundation 2018/0410; The Governmental Funding of Clinical Research within the National Health Service (ALF) 2018/40304; The Swedish Cancer Society 2019/0388; The Swedish Research Council 2020/01491; The Swedish Cancer- and Allergy Fund, ALF young researcher Looket Dihge; ALF young researcher Ida Skarping. The funders had no role in the study design; collection, analyses, or interpretation of data; writing of the manuscript; or decision to publish the results.

References

- Sung H, Ferlay J, Siegel RL, Laversanne M, Soerjomataram I, Jemal A, et al. Global cancer statistics 2020: Globocan estimates of incidence and mortality worldwide for 36 cancers in 185 countries. *CA Cancer J Clin* (2021) 71(3):209–49. doi: 10.3322/caac.21660
- Keelan S, Flanagan M, Hill ADK. Evolving trends in surgical management of breast cancer: An analysis of 30 years of practice changing papers. *Front Oncol* (2021) 11:622621. doi: 10.3389/fonc.2021.622621
- Carter CL, Allen C, Henson DE. Relation of tumor size, lymph node status, and survival in 24,740 breast cancer cases. *Cancer* (1989) 63(1):181–7. doi: 10.1002/1097-0142(19890101)63:1<181::AID-CNCR2820630129>3.0.CO;2-H
- Veronesi U, Paganelli G, Viale G, Luini A, Zurrada S, Galimberti V, et al. A randomized comparison of sentinel-node biopsy with routine axillary dissection in breast cancer. *New Engl J Med* (2003) 349(6):546–53. doi: 10.1056/NEJMoa012782
- Mansel RE, Fallowfield L, Kissin M, Goyal A, Newcombe RG, Dixon JM, et al. Randomized multicenter trial of sentinel node biopsy versus standard axillary treatment in operable breast cancer: The almanac trial. *J Natl Cancer Inst* (2006) 98(9):599–609. doi: 10.1093/jnci/djj158
- Krag DN, Anderson SJ, Julian TB, Brown AM, Harlow SP, Costantino JP, et al. Sentinel-Lymph-Node resection compared with conventional axillary-Lymph-Node dissection in clinically node-negative patients with breast cancer: Overall survival findings from the nsabp b-32 randomised phase 3 trial. *Lancet Oncol* (2010) 11(10):927–33. doi: 10.1016/S1470-2045(10)70207-2
- Giuliano AE, Hunt KK, Ballman KV, Beitsch PD, Whitworth PW, Blumencranz PW, et al. Axillary dissection vs no axillary dissection in patients with invasive breast cancer and sentinel node metastasis: A randomized clinical trial. *Jama* (2011) 305(6):569–75. doi: 10.1001/jama.2011.90
- Galimberti V, Cole BF, Zurrada S, Viale G, Luini A, Veronesi P, et al. Axillary dissection versus no axillary dissection in patients with sentinel-node micrometastases (Ibcsg 23-01): A phase 3 randomised controlled trial. *Lancet Oncol* (2013) 14(4):297–305. doi: 10.1016/S1470-2045(13)70035-4

Acknowledgments

A preliminary version of this manuscript has been posted as preprint available at Research Square [<https://doi.org/10.21203/rs.3.rs-1475276/v1>] (60).

Conflict of interest

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher's note

All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

Supplementary material

The Supplementary Material for this article can be found online at: <https://www.frontiersin.org/articles/10.3389/fonc.2023.1102254/full#supplementary-material>

- Donker M, van Tienhoven G, Straver ME, Meijnen P, van de Velde CJ, Mansel RE, et al. Radiotherapy or surgery of the axilla after a positive sentinel node in breast cancer (EORTC 10981-22023 amaro): A randomised, multicentre, open-label, phase 3 non-inferiority trial. *Lancet Oncol* (2014) 15(12):1303–10. doi: 10.1016/S1470-2045(14)70460-7
- Giuliano AE, Ballman KV, McCall L, Beitsch PD, Brennan MB, Kelemen PR, et al. Effect of axillary dissection vs no axillary dissection on 10-year overall survival among women with invasive breast cancer and sentinel node metastasis: The acosog Z0011 (Alliance) randomized clinical trial. *Jama* (2017) 318(10):918–26. doi: 10.1001/jama.2017.11470
- Verdial FC, Etzioni R, Duggan C, Anderson BO. Demographic changes in breast cancer incidence, stage at diagnosis and age associated with population-based mammographic screening. *J Surg Oncol* (2017) 115(5):517–22. doi: 10.1002/jso.24579
- Ashikaga T, Krag DN, Land SR, Julian TB, Anderson SJ, Brown AM, et al. Morbidity results from the nsabp b-32 trial comparing sentinel lymph node dissection versus axillary dissection. *J Surg Oncol* (2010) 102(2):111–8. doi: 10.1002/jso.21535
- Crane-Okada R, Wascher RA, Elashoff D, Giuliano AE. Long-term morbidity of sentinel node biopsy versus complete axillary dissection for unilateral breast cancer. *Ann Surg Oncol* (2008) 15(7):1996–2005. doi: 10.1245/s10434-008-9909-y
- Schulze T, Mucke R, Markwardt J, Schlag PM, Bembek A. Long-term morbidity of patients with early breast cancer after sentinel lymph node biopsy compared to axillary lymph node dissection. *J Surg Oncol* (2006) 93(2):109–19. doi: 10.1002/jso.20406
- Kootstra JJ, Dijkstra PU, Rietman H, de Vries J, Baas P, Geertzen JHB, et al. A longitudinal study of shoulder and arm morbidity in breast cancer survivors 7 years after sentinel lymph node biopsy or axillary lymph node dissection. *Breast Cancer Res Treat* (2013) 139(1):125–34. doi: 10.1007/s10549-013-2509-y
- Langer I, Guller U, Berclaz G, Koechli OR, Schaefer G, Fehr MK, et al. Morbidity of sentinel lymph node biopsy (SlN) alone versus sln and completion axillary lymph

- node dissection after breast cancer surgery - a prospective Swiss multicenter study on 659 patients. *Ann Surg* (2007) 245(3):452–61. doi: 10.1097/01.sla.0000245472.47748.ec
17. Brackstone M, Baldassarre FG, Perera FE, Cil T, Chavez Mac Gregor M, Dayes IS, et al. Management of the axilla in early-stage breast cancer: Ontario health (Cancer care Ontario) and asco guideline. *J Clin Oncol* (2021) 39(27):3056–82. doi: 10.1200/jco.21.00934
18. van Wely BJ, de Wilt JH, Francissen 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(3):159–68. doi: 10.1002/bjs.9663
19. Kamitani T, Hatakenaka M, Yabuuchi H, Matsuo Y, Fujita N, Jinnouchi M, et al. Detection of axillary node metastasis using diffusion-weighted mri in breast cancer. *Clin Imaging* (2013) 37(1):56–61. doi: 10.1016/j.clinimag.2012.02.014
20. Kuijs VJ, Moosdorff M, Schipper RJ, Beets-Tan RG, Heuts EM, Keymeulen KB, et al. The role of mri in axillary lymph node imaging in breast cancer patients: A systematic review. *Insights Imaging* (2015) 6(2):203–15. doi: 10.1007/s13244-015-0404-2
21. Lisboa PJ, Taktak AFG. The use of artificial neural networks in decision support in cancer: A systematic review. *Neural Networks* (2006) 19(4):408–15. doi: 10.1016/j.neunet.2005.10.007
22. Sargent DJ. Comparison of artificial neural networks with other statistical approaches: Results from medical data sets. *Cancer* (2001) 91(8 Suppl):1636–42. doi: 10.1002/1097-0142(20010415)91:8+<1636::AID-CNCR1176>3.0.CO;2-D
23. Dihge L, Ohlsson M, Eden P, Bendahl PO, Ryden L. Artificial neural network models to predict nodal status in clinically node-negative breast cancer. *BMC Cancer* (2019) 19(1):610. doi: 10.1186/s12885-019-5827-6
24. Skarping I, Nilsson K, Dihge L, Fridhammar A, Ohlsson M, Huss L, et al. The implementation of a noninvasive lymph node staging (Nils) preoperative prediction model is cost effective in primary breast cancer. *Breast Cancer Res Treat* (2022) 194(3):577–86. doi: 10.1007/s10549-022-06636-x
25. Krogh A, Hertz JA. (1991)., in: *Proceedings of the 4th International Conference on Neural Information Processing Systems (Nips'91)*, (San Francisco, CA, USA: Morgan Kaufmann Publishers Inc).
26. Luo W, Phung D, Tran T, Gupta S, Rana S, Karmakar C, et al. Guidelines for developing and reporting machine learning predictive models in biomedical research: A multidisciplinary view. *J Med Internet Res* (2016) 18(12). doi: 10.2196/jmir.5870
27. Trevena LJ, Zikmund-Fisher BJ, Edwards A, Gaissmaier W, Galesic M, Han PKJ, et al. Presenting quantitative information about decision outcomes: A risk communication primer for patient decision aid developers. *BMC Med Inf Decision Making* (2013) 13(S2). doi: 10.1186/1472-6947-13-s2-s7
28. Spiegelhalter D. Risk and uncertainty communication. *Annu Rev Stat Its Appl* (2017) 4(1):31–60. doi: 10.1146/annurev-statistics-010814-020148
29. Skarping I, Dihge L, Bendahl P-O, Huss L, Ellbrant J, Ohlsson M, et al. The Nils study protocol: A retrospective validation study of an artificial neural network based preoperative decision-making tool for noninvasive lymph node staging in women with primary breast cancer (Isrctn14341750). *Diagnostics* (2022) 12(3):582. doi: 10.3390/diagnostics12030582
30. Sharifi S, Peterson MK, Baum JK, Raza S, Schnitt SJ. Assessment of pathologic prognostic factors in breast core needle biopsies. *Mod Pathol* (1999) 12(10):941–5.
31. Harris GC, Denley HE, Pinder SE, Lee AH, Ellis IO, Elston CW, et al. Correlation of histologic prognostic factors in core biopsies and therapeutic excisions of invasive breast carcinoma. *Am J Surg Pathol* (2003) 27(1):11–5. doi: 10.1097/0000478-200301000-00002
32. Wood B, Junckerstorff R, Sterrett G, Frost F, Harvey J, Robbins P. A comparison of immunohistochemical staining for oestrogen receptor, progesterone receptor and her-2 in breast core biopsies and subsequent excisions. *Pathology* (2007) 39(4):391–5. doi: 10.1080/00313020701444445
33. Romero Q, Bendahl PO, Klinton M, Loman N, Ingvar C, Ryden L, et al. Ki67 proliferation in core biopsies versus surgical samples - a model for neo-adjuvant breast cancer studies. *BMC Cancer* (2011) 11:341. doi: 10.1186/1471-2407-11-341
34. Tawfik K, Kimler BF, Davis MK, Fan F, Tawfik O. Ki-67 expression in axillary lymph node metastases in breast cancer is prognostically significant. *Hum Pathol* (2013) 44(1):39–46. doi: 10.1016/j.humpath.2012.05.007
35. Bevilacqua JL, Kattan MW, Fey JV, Cody HS3rd, 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 Off J Am Soc Clin Oncol* (2007) 25(24):3670–9. doi: 10.1200/JCO.2006.08.8013
36. 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(3):839–48. doi: 10.1007/s10549-012-2219-x
37. Mojarad S, Venturini B, Fulgenzi P, Papaleo R, Brisigotti M, Monti F, et al. Prediction of nodal metastasis and prognosis of breast cancer by Ann-based assessment of tumour size and P53, ki-67 and steroid receptor expression. *Anticancer Res* (2013) 33(9):3925–33.
38. Mattfeldt T, Kestler HA, Sinn HP. Prediction of the axillary lymph node status in mammary cancer on the basis of clinicopathological data and flow cytometry. *Med Biol Eng Comput* (2004) 42(6):733–9. doi: 10.1007/BF02345205
39. Zhou L-Q, Wu X-L, Huang S-Y, Wu G-G, Ye H-R, Wei Q, et al. Lymph node metastasis prediction from primary breast cancer us images using deep learning. *Radiology* (2020) 294(1):19–28. doi: 10.1148/radiol.2019190372
40. Zheng X, Yao Z, Huang Y, Yu Y, Wang Y, Liu Y, et al. Deep learning radiomics can predict axillary lymph node status in early-stage breast cancer. *Nat Commun* (2020) 11(1). doi: 10.1038/s41467-020-15027-z
41. Zhang Z, Beck MW, Winkler DA, Huang B, Sibanda W, Goyal H, et al. Opening the black box of neural networks: Methods for interpreting neural network models in clinical applications. *Ann Transl Med* (2018) 6(11):216. doi: 10.21037/atm.2018.05.32
42. Cortez P, Embrechts MJ. Using sensitivity analysis and visualization techniques to open black box data mining models. *Inf Sci* (2013) 225:1–17. doi: 10.1016/j.ins.2012.10.039
43. Mazo C, Kearns C, Mooney C, Gallagher WM. Clinical decision support systems to open black box data mining models. *Inf Sci* (2013) 225:1–17. doi: 10.1016/j.ins.2012.10.039
44. Ellis PG, Brufsky AM, Beriwal S, Lokay KG, Benson HO, McCutcheon SB, et al. Pathways clinical decision support for appropriate use of key biomarkers. *J Oncol Pract* (2016) 12(6):e681–7. doi: 10.1200/JOP.2015.010546
45. Wishart GC, Azzato EM, Greenberg DC, Rashbass J, Kearns O, Lawrence G, et al. Predict: A new uk prognostic model that predicts survival following surgery for invasive breast cancer. *Breast Cancer Res* (2010) 12(1):R1. doi: 10.1186/bcr2464
46. Sutton RT, Pincock D, Baumgart DC, Sadowski DC, Fedorak RN, Kroeker KI. An overview of clinical decision support systems: Benefits, risks, and strategies for success. *NPJ Digit Med* (2020) 3:17. doi: 10.1038/s41746-020-0221-y
47. Segal MM, Rahm AK, Hulse NC, Wood G, Williams JL, Feldman L, et al. Experience with integrating diagnostic decision support software with electronic health records: Benefits versus risks of information sharing. *EGEMS (Wash DC)* (2017) 5(1):23. doi: 10.5334/egems.244
48. *Choosing wisely: Society of surgical oncology* (2016). Available at: <https://www.choosingwisely.org/clinician-lists/sso-sentinel-node-biopsy-in-node-negative-women-70-and-over/>.
49. McKeivitt E, Cheifetz R, DeVries K, Laws A, Warburton R, Gondara L, et al. Sentinel node biopsy should not be routine in older patients with er-positive Her2-negative breast cancer who are willing and able to take hormone therapy. *Ann Surg Oncol* (2021) 28(11):5950–7. doi: 10.1245/s10434-021-09839-6
50. Gentilini 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(5):678–81. doi: 10.1016/j.breast.2012.06.013
51. 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/It1-2) in the context of breast conserving therapy: First results following commencement of the intergroup-Sentinel-Mamma (Insema) trial. *Geburtsh Frauenheilk* (2017) 77(2):149–57. doi: 10.1055/s-0042-122853
52. Keelan S, Heeney A, Downey E, Hegarty A, Roche T, Power C, et al. Breast cancer patients with a negative axillary ultrasound may have clinically significant nodal metastasis. *Breast Cancer Res Treat* (2021) 187(2):303–10. doi: 10.1007/s10549-021-06194-8
53. Collins GS, Reitsma JB, Altman DG, Moons KG. Transparent reporting of a multivariable prediction model for individual prognosis or diagnosis (Tripod): The tripod statement. *BMJ* (2015) 350:g7594. doi: 10.1136/bmj.g7594
54. Lagerlund M, Åkesson A, Zackrisson S. Population-based mammography screening attendance in Sweden 2017–2018: A cross-sectional register study to assess the impact of sociodemographic factors. *Breast* (2021) 59:16–26. doi: 10.1016/j.breast.2021.05.011
55. Dawson SJ, Duffy SW, Blows FM, Driver KE, Provenzano E, LeQuesne J, et al. Molecular characteristics of screen-detected vs symptomatic breast cancers and their impact on survival. *Br J Cancer* (2009) 101(8):1338–44. doi: 10.1038/sj.bjc.6605317
56. Saerens M, Latinne P, Decaestecker C. Adjusting the outputs of a classifier to new a priori probabilities: A simple procedure. *Neural Comput* (2002) 14(1):21–41. doi: 10.1162/089976602753284446
57. Sola M, Alberro JA, Fraile M, Santesteban P, Ramos M, Fabregas R, et al. Complete axillary lymph node dissection versus clinical follow-up in breast cancer patients with sentinel node micrometastasis: Final results from the multicenter clinical trial aatrm 048/13/2000. *Ann Surg Oncol* (2013) 20(1):120–7. doi: 10.1245/s10434-012-2569-y
58. Krop I, 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(24):2838–47. doi: 10.1200/JCO.2017.74.0472
59. Cardoso F, Kyriakides S, Ohno S, Penault-Llorca F, Poortmans P, Rubio IT, et al. Early breast cancer: Esmo clinical practice guidelines for diagnosis, treatment and follow-up. *Ann Oncol* (2019) 30(10):1674. doi: 10.1093/annonc/mdz189
60. Dihge L, Bendahl P-O, Skarping I, Hjårtström M, Ohlsson M, Ryden L. Nils: An artificial neural network tool for personalized noninvasive lymph node staging in breast cancer. *Res Square* (2022). doi: 10.21203/rs.3.rs-1475276/v1