A Comparison of Breast Surgeon and Nomogram-Generated Risk Predictions of Sentinel and Non-Sentinel Node Metastases*

Luisa Sugaya¹, Paulo R. de Alcantara Filho¹,², Bruna Salani Mota¹, Sujata Patil³, Kimberly J. Van Zee², José Luiz B. Bevilacqua¹#

¹Hospital Sirio-Libanes, Sao Paulo, Brazil; ²Breast Service, Department of Surgery, Memorial Sloan-Kettering Cancer Center, New York, USA; ³Department of Epidemiology and Biostatistics, Memorial Sloan-Kettering Cancer Center, New York, USA.
Email: ²bevilacqua@hsl.org.br

Received June 3rd, 2013; revised July 2nd, 2013; accepted July 9th, 2013

Copyright © 2013 Luisa Sugaya et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

ABSTRACT
Memorial Sloan-Kettering Cancer Center (MSKCC) has developed 2 nomograms: the Sentinel Lymph Node Nomogram (SLNN), which is used to predict the likelihood of sentinel lymph node (SLN) metastases in patients with invasive breast cancer, and the Non-Sentinel Lymph Node Nomogram (NSLNN), which is used to predict the likelihood of residual axillary disease after a positive SLN biopsy. Our purpose was to compare the accuracy of MSKCC nomogram predictions with those made by breast surgeons. Two questionnaires were built with characteristics of two sets of 33 randomly selected patients from the MSKCC Sentinel Node Database. The first included only patients with invasive breast cancer, and the second included only patients with invasive breast cancer and positive SLN biopsy. 26 randomly selected Brazilian breast surgeons were asked about the probability of each patient in the first set having SLN metastases and each patient in the second set having additional non-SLN metastases. The predictions of the nomograms and breast surgeons were compared. There was no correlation between nomogram risk predictions and breast surgeon risk prediction estimates for either the SLNN or the NSLNN. The area under the receiver operating characteristics curves (AUCs) were 0.871 and 0.657 for SLNN and breast surgeons, respectively (p < 0.0001), and 0.889 and 0.575 for the NSLNN and breast surgeons, respectively (p < 0.0001). The nomograms were significantly more accurate as prediction tools than the risk predictions of breast surgeons in Brazil. This study demonstrates the potential utility of both nomograms in the decision-making process for patients with invasive breast cancer.

Keywords: Sentinel Lymph Node Biopsy; Nomogram; Predictions; Breast Cancer; Completion Axillary Lymph Node Dissection

1. Introduction

Over the last decades, public awareness, medical education, and increased use of screening mammography have resulted in earlier detection and treatment of invasive breast cancer, greatly improving the prognosis of those patients [1,2]. The adoption of sentinel lymph node biopsy (SLNB) has allowed those with no axillary metastases to avoid axillary lymph node dissection (ALND), and has reduced morbidity as a result. Since the adoption of SLNB to stage the axilla, the standard management of a positive SLNB has included ALND. However, only 40% - 50% of patients with positive sentinel lymph nodes (SLNs) who undergo completion axillary lymph node dissection (cALND) have additional non-SLN metastases, with favorable subsets of women having an even lower risk of non-SLN metastases. Women without residual nodal disease are unlikely to benefit from cALND, but are exposed to the risk of lymphedema and other morbidities.

A recent randomized controlled trial of cALND (American College of Surgeons Oncology Group [ACOSOG] Z0011) was carried out in a selected group of clinically node-negative women with early-stage breast cancer undergoing breast-conserving surgery and whole breast radiation, with positive SLN. With over 6 years of follow-
up, no statistically significant differences in local or regional recurrence were found in the groups that had SLNB alone and those that underwent cALND [3]. These findings support previous studies that suggested there is a subgroup of patients with favorable characteristics who could avoid ALND and be treated with breast-conserving surgery, radiotherapy, and adjuvant systemic therapy [4-9].

A recent study from Memorial Sloan-Kettering Cancer Center (MSKCC) analyzed the possible impact of Z0011 on the rate of ALND using the criteria of that trial [10]. The study concluded that if the Z0011 trial criteria were applied to the MSKCC SLN database population, the need for ALND would have been reduced by 13% in the overall breast cancer patient population and by 48% in the subset of patients with positive SLNs. In fact, in the United States, many of those with SLN micrometastases have not been undergoing cALND; an analysis of the National Cancer Data Base showed that in 2005 almost 50% of those with microscopic metastases in the SLN did not undergo ALND [11].

In order to estimate the likelihood of SLN metastases or to select women in whom the risk of having residual axillary disease is low, it is necessary to weigh multiple, different factors which have been shown to be predictive: tumor; sentinel node; and patient characteristics. It is difficult for a clinician to appropriately weigh all of the various factors, which increase or decrease the risk of disease simultaneously. Nomograms are useful tools that improve the accuracy of clinical judgment [12-16] and help clinicians to deal with these very complex tasks. Nomograms are prediction instruments based on regression analysis of multiple variables that translate complex calculations to simple graphical tools or simple software. In addition, nomograms also permit clinicians to provide more precise and comprehensible information to patients, who are increasingly interested in their disease and who are increasingly taking part in the decisions regarding their treatment [17].

MSKCC has built 2 prognostic nomograms, one to predict the likelihood of SLN metastases in patients with invasive breast cancer (the MSKCC SLN nomogram) [18], and one to predict the likelihood of non-SLN metastases in patients with a positive SLN biopsy (the MSKCC Non-SLN nomogram) [19]. Both are available from the MSKCC website at http://www.mskcc.org/cancer-care/adult/breast/prediction-tools-and-at www.nomograms.org. The MSKCC SLN nomogram (SLNN) incorporates 9 characteristics of primary breast cancer, including age, tumor size, tumor type, lymphovascular invasion, tumor location, multifocality, nuclear grade, and presence of hormone receptors. The MSKCC non-SLN nomogram (NSLNN) includes the same variables, except for age and tumor location, plus characteristics of the SLNs.

The NSLNN has been validated worldwide, showing its robustness in populations different from the one in which it was developed [20-24]. However, a nomogram’s utility is determined by its superiority over clinical judgment [25]. In contrast to use in the United States and Europe, the use of nomograms in Brazil and other developing countries is limited. Here we seek to compare the risk predictions of SLN metastases generated by both nomograms with risk predictions of SLN metastases based on the clinical judgment of breast surgeons in Brazil.

2. Materials and Methods

This project was approved by the Institutional Review Boards of Hospital SirioLibanês, São Paulo, Brazil (HSL2008/13) and MSKCC. The NSLNN and SLNN were each developed in a training population, and each was then tested in an independent validation population from MSKCC [18,19]. For the present study, all patient data were de-identified. We randomly selected 33 cases from each of the validation populations used to test the original nomograms.

Of 166 breast surgeons registered with the Brazilian Society of Breast Surgeons and working in São Paulo, we randomly selected 81 participants. Twenty-six breast surgeons agreed to participate; the others either declined to participate (n = 7) or were unable to be contacted (n = 48). All participants were volunteers and provided signed informed consent. Most of these breast surgeons were unfamiliar with the nomograms, and none had incorporated their use into their clinical practice at the time of the study.

Two questionnaires were administered to each surgeon. In the first, data from 33 patients randomly selected from the SLNN validation population were provided: age; tumor size; tumor type; presence of lymphovascular invasion; tumor location; multifocality; nuclear grade; and estrogen and progesterone receptor status. For each patient, breast surgeons were asked: “Of 100 women with invasive breast cancer with these characteristics, how many of them would you expect to have one or more positive sentinel lymph nodes?”

In the second questionnaire, data were presented from 33 cases from the NSLNN validation population: pathological size; tumor type; nuclear grade; lymphovascular invasion; multifocality and estrogen-receptor status of the primary tumor; method of detection of SLN metastases; number of positive SLNs; and number of negative SLNs. Breast surgeons were asked: “If 100 women with a positive sentinel node had these characteristics and then received a full axillary dissection, how many of them would you expect to have one or more positive non-sentinel lymph nodes?”

When a response was provided as a range, the highest
value was used. Breast surgeon risk predictions were plotted against MSKCC nomogram risk predictions to assess correlation. To compare the accuracy of each nomogram’s risk predictions with those of the breast surgeons, we calculated the area under the receiver operating curve (AUC-ROC). AUC estimates and standard errors were calculated using bootstrap methods to account for each patient having been evaluated once by each surgeon.

To compare the AUCs for each nomogram and from the surgeon estimates, we used bootstrap methods. Specifically, we bootstrapped each of the 33 scenarios 200 times and calculated the AUCs for the nomogram risk predictions and surgeon risk predictions for each bootstrapped dataset. The differences were tested using a paired t-test.

3. Results

Twenty-six breast surgeons, 32.1% of our initial sample, answered the 2 questionnaires completely. Histograms of the nomogram predictions and surgeon estimates by actual outcomes revealed different distributions for the nomograms and breast surgeons (Figures 1 and 2). In both scenarios, the nomogram predictions better differentiated negative and positive outcomes, particularly when predicting the presence of residual axillary diseases after a positive SLN biopsy (Figure 2). The predictions of breast surgeons were similar for cases with positive and negative outcomes.

When predicting the likelihood of a patient having SLN metastases, the AUC-ROC achieved by the MSKCC nomograms was 0.871 (Figure 3). In comparison, the AUC-ROC achieved by the breast surgeons in our study was 0.657 (p < 0.0001) (Figure 3). When predicting the likelihood of a patient having non-SLN metastases when their SLN was positive, the AUC-ROC achieved by the nomogram was 0.889 (Figure 3); whereas the AUC-ROC achieved by clinical experts was 0.575 (p < 0.0001) (Figure 3).

4. Discussion

To our knowledge, this is the first study to compare risk predictions of SLN metastases made by clinical experts to
those generated by MSKCC’s 2 nomograms. Our findings are in agreement with previous studies [15,25,26] suggesting that nomograms outperform clinicians in predicting clinical outcomes related to multiple variables. And, similar to prior studies [25,26], we found that the NSLNN was more accurate than the breast surgeons in our study in predicting the likelihood of non-SLN metastases in patients with positive SLN biopsy. Our findings support the utility of nomograms in predicting these outcomes.

Several factors may explain the superior accuracy of the MSKCC nomograms. The use of regression models ensures that each variable contributes to the results based on their actual predictive power and relation to the criterion of interest. Therefore, if using the same data, nomograms always lead to the same conclusion, preventing the influence of random factors that often bias clinical judgment [27]. Furthermore, for clinicians, it is difficult to clearly distinguish between valid and invalid variables, and false beliefs often develop regarding associations between variables. Recall bias, control bias, fatigue, and clinical experience can also influence clinical judgment [28]. Nomograms translate complex equations through simple graphics or software, and are powerful tools that help clinicians and patients balance benefits and risks, while managing the diagnosis and treatment of the disease.

The recently published ACOSOG Z0011 randomized controlled trial of axillary node dissection in a selected group of women with early-stage breast cancer and positive SLN undergoing whole breast radiation and systemic adjuvant treatment suggested that in women with a low risk of further metastasis, ALND may have no therapeutic benefit [3] and results in increased morbidity [29-34].

During the last decade, after the introduction of the NSLNN, there has been a declining rate of ALND in patients with positive SLN [11,19], and several groups have reported a very low rate (<1%) of axillary recurrence among patients selected to not undergo ALND [6,35-37]. At MSKCC, when compared to those patients who did undergo cALND after a positive SLN, patients who did not undergo ALND were older, more likely to have breast conservation, and had smaller tumors, fewer high-grade lesions, less lymphovascular invasion, and lower nomogram scores overall [8], suggesting that the nomogram did identify a group at low risk of axillary recurrence.

The universal use of SLN biopsy is less controversial than that of cALND, and therefore, the SLNN would not often result in a change in surgical decision making. However, the SLNN nomogram helps a breast surgeon explain to a woman the risk of nodal metastasis of her disease. The absolute risk prediction provided by the nomogram is much more useful than knowing a list of risk factors, each associated with a hazard ratio, and can assist in patient education and decision making.

Nevertheless, it is clear that nomograms cannot replace clinical judgment. In deciding whether or not to perform SLN biopsy or completion ALND, breast surgeons also need to consider inaccuracies in the diagnoses, uncertainties about the progression of the disease, variable responses to treatment, implications for systemic therapy, and the individual characteristics of each patient, such as age and psychological factors [38].

With the results of ACOSOG Z0011, the increasing use of nomograms, the development of more accurate tools [39,40], and a clearer identification of a subgroup of patients who can safely avoid such procedures, breast surgeons in Brazil may be encouraged to include the nomograms in their decision-making processes after discussing both related advantages and disadvantages, and the impact of avoiding the completion of ALND.

One may argue that the major limitation of the recent randomized Z0011 trial is that of low external validity because that trial selected patients at very low risk of additional metastases. Indeed, the authors of that trial state that “most of the patients in this trial had a low axillary tumor burden” [3]. Caution at the initiation of the study led to an attempt to assure that women with high tumor burden were not randomized to SLND alone. The Z0011 trial excluded patients who underwent mastectomy, had clinically palpable lymph nodes, had 3 or more positive SLNs, had large tumors, and who received neoadjuvant treatment. As a result of these careful selection criteria, in the women randomized to ALND, only 27% had additional nodal disease identified. This is significantly lower than the more typical rate of 40% - 50% in most unselected series of clinically node-negative women who are found to have a positive SLN [19,39]. According to NCCN guidelines [41], completion ALND is still a standard of care for positive-SLN patients, especially in those patients who do not fulfill the criteria of ACOSOG Z0011 (women with larger tumors; women with clinically palpable nodes or 3 or more positive SLNs; women undergoing mastectomy or lumpectomy without whole breast radiation). For this population, the nomogram is a useful tool to predict additional metastasis.

Although all the data in this study that were presented to the breast surgeons were taken from real cases, clinical decisions based on hypothetical situations may not reproduce real clinical scenarios. Other variables, such as years of practice, the type of hospital that the breast surgeon interviewed worked at, previous contact with other nomograms, and willingness to participate in the study may also have influenced the results. The AUC-ROC of both nomograms were larger than the ones found in the original studies [18,19], and we randomly selected the
cases for our questionnaires from patients that were not included in the nomogram development populations.

In order to achieve more individualized treatment of patients with breast cancer, breast surgeons will need to accurately predict the probability of a patient having additional disease. Our study demonstrates that the risk predictions of both nomograms were more precise than the risk predictions estimated by breast surgeons, demonstrating the potential utility of both nomograms in the decision-making process for patients with invasive breast cancer.

REFERENCES


