The Breast 22 (2013) 869–873

Contents lists available at SciVerse ScienceDirect

The Breast

journal homepage: www.elsevier.com/brst

Original article

A nomogram for predicting underestimation of invasiveness in ductal carcinoma in situ diagnosed by preoperative needle biopsy

Hyung Seok Park [a], Ha Yan Kim [b], Seho Park [a], Eun-Kyung Kim [c], Seung Il Kim [a], Byeong-Woo Park [a,d,*]

[a] Department of Surgery, Yonsei University College of Medicine, 50 Yonsei-ro, Seodaemun-gu, Seoul 120-752, Republic of Korea
[b] Biostatistics Collaboration Unit, Yonsei University College of Medicine, Seoul, Republic of Korea
[c] Department of Radiology, Yonsei University College of Medicine, Seoul, Republic of Korea
[d] Brain Korea 21 Project for Medical Science, Yonsei University College of Medicine, Seoul, Republic of Korea

ARTICLE INFO

Article history:
Received 11 December 2012
Received in revised form
21 February 2013
Accepted 19 March 2013

Keywords:
Noninfiltrating intraductal carcinoma
Needle biopsy
Mammography
Mammary ultrasonography
Nomogram
Axillary metastasis

ABSTRACT

It is unnecessary to perform axillary staging in patients with ductal carcinoma in situ (DCIS) of the breast because of the low incidence of axillary metastasis. However, diagnosis of DCIS by core needle biopsy showed a high rate of underestimation of invasive cancer. Thus, it is necessary to predict invasiveness in DCIS patients on core before surgery. We analyzed 340 patients with DCIS diagnosed by needle biopsy. The cases were divided into training and validation sets. Logistic regression was performed to predict the presence of invasive cancer in the final pathology, and a nomogram was constructed from the training set using the presence of palpability, the presence of ultrasonographic calcification and mass, the biopsy tools, and the presence of microinvasion. The model was subsequently applied to the validation set. The nomogram for the training set was both accurate and discriminating, with an area under the receiver operating characteristic curve (AUC) of 0.75. When applied to the validation group, the model accurately predicted the likelihood of invasive cancer (AUC: 0.71). Our nomogram will allow surgeons to easily and accurately estimate the likelihood of invasive cancer in patients with DCIS as diagnosed by preoperative needle biopsy.

Introduction

Underestimation of invasive cancer in patients diagnosed with ductal carcinoma in situ (DCIS) by preoperative needle biopsy has led to some debate over whether to perform sentinel lymph node biopsy (SLNB) on patients at the time of surgery.[1,2] The low incidence of axillary metastasis in DCIS patients supports the omission of SLNB,[3] and thus national comprehensive cancer network guidelines do not recommend SLNB for patients with DCIS.[4] On the other hand, the high rate of underestimation of invasive cancer in DCIS patients who were diagnosed by core needle biopsy provides justification for performing SLNB at the time of surgery.[5] Numerous studies have analyzed the risk factors for underestimation of invasive cancer in preoperatively diagnosed DCIS patients.[6–13] These studies suggested that histologic grade, biopsy methods, palpability, mammographic or sonographic findings, and presence of microinvasive focus were associated with underestimation; however, none of these factors have been accepted as a definitive predictor of underestimation. In an attempt to develop a more precise prediction method than the identification of risk predictors for underestimation, a previous study demonstrated that the number of predictors was associated with the probability of invasive cancer in patients with DCIS.[14] However, this method is difficult to apply in clinical practice for several reasons: studies performed in different institutions often identified different predictors, there was no clear cut-off value for the number of predictors in the model when doing a crude counting of predictors, and a lack of weighting of the risk predictors could compromise the prediction of underestimation of DCIS. Therefore, a more comprehensive statistical analysis for predicting underestimation in patients with DCIS is necessary.

With this aim, we evaluated the risk factors that are associated with underestimation and developed a nomogram that can determine the probability of underestimation in individual DCIS patients who were diagnosed by needle biopsy.

Patients and methods

A nomogram was developed based on patient records from the breast cancer database of Yonsei University Severance Hospital

* Corresponding author. Department of Surgery, Yonsei University College of Medicine, 50 Yonsei-ro, Seodaemun-gu, Seoul 120-752, Republic of Korea. Tel.: +82 2 2228 2100; fax: +82 2 313 8289.
E-mail address: bwpark@yuhs.ac (B.-W. Park).

0960-9776/$ – see front matter © 2013 Elsevier Ltd. All rights reserved.
http://dx.doi.org/10.1016/j.breast.2013.03.009

from 2000 through 2008. Cases diagnosed by open biopsy were excluded. A total of 340 patients who were preoperatively diagnosed with DCIS and underwent definitive surgery were initially included in this study. After exclusion of 10 patients with missing variables for significant predictors, a total of 330 patients were used to establish the nomogram.

The breast cancer database of Yonsei University Severance Hospital is programmed using MS Access (Microsoft, USA) and contains patients' clinical characteristics, pathologic data from preoperative or postoperative evaluations, treatment methods, recurrence data, preoperative evaluation findings, including findings from physical examinations, mammography, and ultrasonography, and follow-up data. The database was analyzed retrospectively. All patients received breast conservation surgery or mastectomy, with or without axillary lymph node staging, including SLNB, axillary lymph node dissection, or both. This study was approved by the Institutional Review Board of Severance Hospital, Yonsei University Health System.

To establish a regression model to develop the nomogram, patients were first divided into two groups according to the final pathology: a pure DCIS group and an invasive cancer group, including microinvasive cancer. Univariate analyses using chi-square tests or Student *t* tests were performed using the results of physical examinations, mammography, ultrasonography, and pathologic findings. Secondly, a binary logistic regression model adjusted for statistically significant factors in the univariate analysis was assessed. Thirdly, the significant factors from a multivariate analysis were used to construct a nomogram. The univariate and multivariate analyses used to identify statistically significant factors were described in detail in our previous study[15]; In brief, the overall underestimation rate of DCIS patients by preoperative needle biopsies was 42.6% (145/340).[15] In univariate analysis, age, operation type, and status of hormone receptors and human epidermal growth factor receptor 2 were not significantly different between patients with pure DCIS and invasive cancer.[15] The DCIS underestimation rate was significantly associated with presence of palpability, biopsy method, mass and calcification by ultrasonography, grade, and the presence of suspicious microinvasion were significantly related to underestimation of invasive cancer in univariate analysis (all $p < 0.05$).[15] Ultrasonographic mass size was marginally significant ($p = 0.057$), however, presence of calcification, mass, asymmetry, distortion, breast image and reporting data system (BI-RADS) category by mammography, BI-RADS category by ultrasonography, presence of comedo necrosis, and Van Nuys grouping of tumors were not significantly associated with underestimation of invasive cancer in the univariate analysis.[15] The multivariate analysis revealed the following significant predictors: the presence of palpability, the presence of an ultrasonographic mass or calcification, biopsy method, and the presence of suspicious microinvasion.[15] DCIS grade was not associated with underestimation of invasive cancer in multivariate analysis ($p = 0.29$).[15] Therefore, we included the five factors including the presence of palpability, the presence of an ultrasonographic mass or calcification, biopsy method, and the presence of suspicious microinvasion as predictors of the nomogram.[15]

To construct and validate the nomogram, we randomly divided all patients into two data sets: a training set and a validation set. The training set contained 70% of all patients (230/330) and the validation set contained 30% of the patients (100/330). The training data were used to establish the nomogram. The discrimination of the model was assessed by using the area under the receiver operating characteristic (ROC) curve. The Hosmer and Lemeshow Goodness-of-fit test was applied to assess whether there was evidence for lack of fit in a logistic regression model. The calibration of the model was assessed graphically and the area under the curve (AUC) was estimated. The model was validated in two ways: internal validation by 200 bootstrapping samples and by applying the validation data set. The internal validations were demonstrated graphically, and the area under the ROC curve was estimated for application of the model to the validation data.

All statistical analyses were performed using SAS 9.2 (SAS Institute Inc., Cary, NC, USA) and R version 2.13.2 (The R Foundation for Statistical Computing, Vienna, Austria) with the design library.

Results

Predictors and basic equations

The estimated beta in training data of the presence of suspicious microinvasion, biopsy method, the presence of ultrasonographic calcification, palpability, and the presence of ultrasonographic mass are shown in Table 1. The absolute value of β was used to measure the rank of the predictors. Points were calculated using the following equation:

$$\text{Points of rank}(n) = [(\text{absolute}(\beta) \text{ of rank}(n)/(\text{absolute}(\beta) \text{ of rank}(n-1)] \times 100$$

For example, points of microinvasion (rank 2) = [(1.36 (absolute (β) of biopsy method (CNB))/(1.80(absolute (β) of rank of suspicious microinvasion)] × 100 = 75.22.

Predicted probability was calculated as follows:

$$\text{Predicted probability} = 1/(1 + \exp(-(\beta_0 + \beta_1 \times X_1 + \beta_2 \times X_2 + \beta_3 \times X_3 + \beta_4 \times X_4 + \beta_5 \times X_5)))$$

For example, the probability of cases with palpability, presence of microinvasion, diagnosis by 14G core needle biopsy (CNB), and presence of ultrasonographic calcification is 92.67%:

$$\text{Predicted probability} = 1/(1 + \exp(2.9045 - 1.8083 \times 1 - 1.3603 \times 1 - 0.9549 \times 1 - 0.6644 \times 1 - 0.6543 \times 1)) = 0.9267$$

Model building and internal validation

A nomogram based on significant predictors was developed using the training data set ($n = 230$) and is illustrated in Fig. 1. A Hosmer and Lemeshow Goodness-of-fit test gave a *p*-value of 0.757. Discrimination of the nomogram was estimated using the area under the ROC curve, which was 0.755. Internal validation was assessed using 200 bootstrapping samples, and yielded a mean absolute error of 0.016 (Fig. 2). Apparent, bias-corrected, and ideal curves are illustrated in Fig. 2. Internal validation was performed by

Table 1
Estimated beta and absolute beta in multivariate analysis of training data ($n = 230$).

Variable	Or (95% CI)	β	Absolute β	Rank	Points
Suspicious microinvasion	6.10 (1.78–20.85)	1.80	1.80	1	100.00
Biopsy method (CNB)	3.89 (1.65–9.15)	1.36	1.36	2	75.22
USG calcification	2.59 (1.39–4.83)	0.95	0.95	3	52.80
Palpability	1.94 (1.02–3.69)	0.66	0.66	4	36.74
USG mass	1.92 (0.89–4.13)	0.65	0.65	5	36.18
Intercept		−2.90			

OR, odds ratio; CI, confidence interval; USG, ultrasonography; CNB, 14G core needle biopsy.